



# STIC Search Report

## Biotech-Chem Library

STIC Database Tracking Number: 114402

TO: Minh-Tam Davis  
Location: REM-3C18/3A24  
Art Unit: 1642  
Monday, February 23, 2004

Case Serial Number: 09899569

From: Toby Port  
Location: Biotech-Chem Library  
Remsen 1A59  
Phone: 571-272-2523

[toby.port@uspto.gov](mailto:toby.port@uspto.gov)

### Search Notes

Dear Examiner Davis,

Here are the results of your search.  
Please feel free to contact me if you have any questions.

Toby Port



***This Page Blank (uspio)***



XX 17-JAN-2002.  
PD 05-JUL-2001; 2001WO-EP07705.  
XX 07-JUL-2000; 2000DE-1031080.  
XX 19-APR-2001; 2001DE-1019294.  
XX (BOEH) BOEHRINGER INGELHEIM INT GMBH.  
PI Schaefer N, Scherl-Mostagier M, Sommergruber W, Abseher R;  
XX MPI, 2002-171704/22.  
XX P-PSDB; AAM49641.  
PT New tumor-associated antigen B345, useful for diagnosis and  
XX immunotherapy of tumors, also related nucleic acid and antibodies -  
XX Claim 3; Page 79-85; 102pp; German.  
XX This invention describes a novel tumour-associated antigen, designated  
CC B345 which has cytostatic activity. B345 is involved in communication,  
CC interaction and/or signal transduction with extracellular components and  
CC ligands, especially in the metastatic potential of cancers, particularly  
CC of the colon. B345 or its immunogenic fragments, also the DNA that  
CC encodes it, are useful for immunotherapy of cancer, particularly  
CC carcinoma of lung or colon. Antibodies raised against B345 are useful for  
CC treatment and diagnosis of cancers that are associated with B345  
CC expression, including their use for targeted delivery of cytotoxic or  
CC radioactive agents. Probes derived from B345 can be used to detect  
CC tumour-specific mutations in the B345 sequence, and can be used to screen  
CC for B345 specific modulators. This sequence encodes a human B345  
CC tumour-associated antigen described in the invention.  
SQ Sequence 6163 BP; 1501 A; 1689 C; 1482 G; 1491 T; 0 other;  
Query Match 100.0%; Score 6163; DB 24; Length 6163;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 6163; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CCACACCGCGAATGGGAGAGTAGAGGAGCCAGCAACCGGTCGCGGAGCCCTGCACCC 60  
DB 1 CCACACCGCGAATGGGAGAGTAGAGGAGCCAGCAACCGGTCGCGGAGCCCTGCACCC 60  
QY 61 TGGGAGGAGAGGCGGTCGCTGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120  
DB 61 TGGGAGGAGAGGCGGTCGCTGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120  
QY 121 CGGGTAGCTCGGAGATCTGCTGCGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180  
DB 121 CGGGTAGCTCGGAGATCTGCTGCGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180  
QY 121 CGGGTAGCTCGGAGATCTGCTGCGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180  
DB 121 CGGGTAGCTCGGAGATCTGCTGCGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180  
QY 181 GTGAGTAGAGCAGAGGAGGAGAGCGAGCTGAGCGCGGGGCTGGGGCGCGCGCTCC 240  
DB 181 GTGAGTAGAGCAGAGGAGGAGAGCGAGCTGAGCGCGGGGCTGGGGCGCGCGCTCC 240  
QY 241 CCACCGTCGTTTTCCCAACCGAGCGCGAGCGTCCCGAGTCAATGAGCGCGCTGAAC 300  
DB 241 CCACCGTCGTTTTCCCAACCGAGCGCGAGCGTCCCGAGTCAATGAGCGCGCTGAAC 300  
QY 301 GGGGTCTCTATCGCACTGCTAGGGGTTCTGCTGCTGGGTGCGGCGCGCTCCGCGG 360  
DB 301 GGGGTCTCTATCGCACTGCTAGGGGTTCTGCTGCTGGGTGCGGCGCGCTCCGCGG 360  
QY 361 GCAGAGCTTTTGAATGCTGCTGCAAGAGAGAAACATTAAGTTCTCATAAAGCTG 420  
DB 361 GCAGAGCTTTTGAATGCTGCTGCAAGAGAGAAACATTAAGTTCTCATAAAGCTG 420  
QY 421 GGGAGCCCGCACTCTGCTGCGAAAACCTGTTACATGCTCATTTCTAAGAACATTAACC 480  
DB 421 GGGAGCCCGCACTCTGCTGCGAAAACCTGTTACATGCTCATTTCTAAGAACATTAACC 480  
QY 481 ATGTTGTCCATCAAGTCTGGAGAAAGAAATAGTCTTAACCTTTAGCTCCAGAGTCTGAG 540

DB 481 ATGTTGTCCATCAAGTCTGGAGAAAGAAATAGTCTTAACCTTTAGCTCCAGAGTCTGAG 540  
QY 541 AATCACTTTGTGATGAGATCCAGAAAAATATGATGTAATGTCAGGCCATGCTCTTT 600  
DB 541 AATCACTTTGTGATGAGATCCAGAAAAATATGATGTAATGTCAGGCCATGCTCTTT 600  
QY 601 GGGGAGGTTCACTTCAAGCCCTCGACATGCTGTTGCTTCACTCCCTCAAGAACTTTATC 660  
DB 601 GGGGAGGTTCACTTCAAGCCCTCGACATGCTGTTGCTTCACTCCCTCAAGAACTTTATC 660  
QY 661 TGGGATGTCAAAGCTCATTAAGAGCATCGGTTTAGAGCTGCAATTTCCATCCCTCGCC 720  
DB 661 TGGGATGTCAAAGCTCATTAAGAGCATCGGTTTAGAGCTGCAATTTCCATCCCTCGCC 720  
QY 721 AGGCAGATGCTCCGGGTAGAGCTGCCAGAGAGTCACTCACTCAACGCGCCGA 780  
DB 721 AGGCAGATGCTCCGGGTAGAGCTGCCAGAGAGTCACTCACTCAACGCGCCGA 780  
QY 781 ATGATGCCACCGTGTGTCAGATGCGAACCTTCTGCAAGATGCACTGTGTCGCGATC 840  
DB 781 ATGATGCCACCGTGTGTCAGATGCGAACCTTCTGCAAGATGCACTGTGTCGCGATC 840  
QY 841 AAGATGCCAAGAGAGTGAATATGCTTACACCTTCCATGCTTCCACCCCAAGATGTC 900  
DB 841 AAGATGCCAAGAGAGTGAATATGCTTACACCTTCCATGCTTCCACCCCAAGATGTC 900  
QY 901 TCCGCTTCAGATTCGCAAAACCGCTCATCTATAAAACGCTGTGATCAATCAAGTCTG 960  
DB 901 TCCGCTTCAGATTCGCAAAACCGCTCATCTATAAAACGCTGTGATCAATCAAGTCTG 960  
QY 961 TTTGAGGAGTGAAGGCTCAAGCAACCTGATGTCGCAACTACCAAGAGGCTTCCCTGAG 1020  
DB 961 TTTGAGGAGTGAAGGCTCAAGCAACCTGATGTCGCAACTACCAAGAGGCTTCCCTGAG 1020  
QY 1021 GATGAGCTTCATGACGTGAGAGTTGTGCTTCTGCAACCTGCGGCGCGCTTCTCTTC 1080  
DB 1021 GATGAGCTTCATGACGTGAGAGTTGTGCTTCTGCAACCTGCGGCGCGCTTCTCTTC 1080  
QY 1081 CTCAACTTCACCTCTCCAACTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140  
DB 1081 CTCAACTTCACCTCTCCAACTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140  
QY 1141 GGGTCCACCAACCAACCCGAGGTGTTAACTGAGAGCAAGACCTGGGAAATGATGCG 1200  
DB 1141 GGGTCCACCAACCAACCCGAGGTGTTAACTGAGAGCAAGACCTGGGAAATGATGCG 1200  
QY 1201 GGGAACTTCAACTCTCTCTGGAAGCTGTGACCAAGATGCCCAAGTCCAGAGATCTTC 1260  
DB 1201 GGGAACTTCAACTCTCTCTGGAAGCTGTGACCAAGATGCCCAAGTCCAGAGATCTTC 1260  
QY 1261 CGGCTGCAAGTTTCAAGTTTGTGTCACATCCAAATGAAGCAATTAATCTTACGTG 1320  
DB 1261 CGGCTGCAAGTTTCAAGTTTGTGTCACATCCAAATGAAGCAATTAATCTTACGTG 1320  
QY 1321 GTTGAATTGAGTAATGAGCGAGCATGTCACTCACATGAGCAACGCGCTCAAAACAG 1380  
DB 1321 GTTGAATTGAGTAATGAGCGAGCATGTCACTCACATGAGCAACGCGCTCAAAACAG 1380  
QY 1381 AGCCGCAAGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440  
DB 1381 AGCCGCAAGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440  
QY 1441 CTCACCTTCAGATCTGAGCTCCAAACAAATCTCTCTTCTTGTGATGATGACAGCT 1500  
DB 1441 CTCACCTTCAGATCTGAGCTCCAAACAAATCTCTCTTCTTGTGATGATGACAGCT 1500  
QY 1501 CTGTGATGAATGATGAGAAAAAACAATAAGCTGACAGCAACCGGTAATGCAAGAGAAA 1560  
DB 1501 CTGTGATGAATGATGAGAAAAAACAATAAGCTGACAGCAACCGGTAATGCAAGAGAAA 1560  
QY 1561 TCTTACTCACTCAGAGTCCAGTGAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620



Db	1561	TCCTACTCACTCCAGGTGTCCTCCAGTCACTTCCACCTGTGTGAGACTGATCTTC	1820
Qy	1621	TCCTGGAAGCTGTGTGTGTCCTCCAGGACAGAGGCTCAGGCTGTGTGTCCAGCCAGAG	1680
Db	1621	TCCTGGAAGCTGTGTGTGTCCTCCAGGACAGAGGCTCAGGCTGTGTGTCCAGCCAGAG	1680
Qy	1661	CTGCAAGCAAGCATCACAGAGAAAGCTCTGCAACACAGCTTCAGTCACTGTGTGCGCACT	1740
Db	1661	CTGCAAGCAAGCATCACAGAGAAAGCTCTGCAACACAGCTTCAGTCACTGTGTGCGCACT	1740
Qy	1741	GGCATACCAGCAGAGGACCTGTACCTTGGGCTCCTTGTGCGCGGGAGGGCTATCAAGAG	1800
Db	1741	GGCATACCAGCAGAGGACCTGTACCTTGGGCTCCTTGTGCGCGGGAGGGCTATCAAGAG	1800
Qy	1801	ATCCAGGTGAAGACAGACATCTCGGTGACCTTGTGCACTTTGTGCCCCAGCTTCCAA	1860
Db	1801	ATCCAGGTGAAGACAGACATCTCGGTGACCTTGTGCACTTTGTGCCCCAGCTTCCAA	1860
Qy	1861	GAGGCTCTCAGGCAAGGCTCTGACGGTGTCTCTTTATCTTATTTTCAAGAGAGCGTT	1920
Db	1861	GAGGCTCTCAGGCAAGGCTCTGACGGTGTCTCTTTATCTTATTTTCAAGAGAGCGTT	1920
Qy	1921	TTCAACGGGACCCCCCGACACAAAAGAGGCTACCTGAGAGACCCCCAATGGAGACCGG	1980
Db	1921	TTCAACGGGACCCCCCGACACAAAAGAGGCTACCTGAGAGACCCCCAATGGAGACCGG	1980
Qy	1981	GGCCTGTCCATCCCTCACCCTCTGTGTCTTGGAACATCAGCGTGTGCCAGAGACAGATGGCC	2040
Db	1981	GGCCTGTCCATCCCTCACCCTCTGTGTCTTGGAACATCAGCGTGTGCCAGAGACAGATGGCC	2040
Qy	2041	TGCTGATCTTTTAAAGAGCGGAGCGGCTGTGTCTGTCCAGACAGGCGGCAATTCA	2100
Db	2041	TGCTGATCTTTTAAAGAGCGGAGCGGCTGTGTCTGTCCAGACAGGCGGCAATTCA	2100
Qy	2101	ATCATTCAGAGACAGCGGACCCCGGCTGAGAGATCTTCACTCTGTGACAGAGATGTGCTC	2160
Db	2101	ATCATTCAGAGACAGCGGACCCCGGCTGAGAGATCTTCACTCTGTGACAGAGATGTGCTC	2160
Qy	2161	CCCAAGCCAACTTCCACCATCACAGCTTCTGTGGGTCAACATCTCTAACTGACGCCACG	2220
Db	2161	CCCAAGCCAACTTCCACCATCACAGCTTCTGTGGGTCAACATCTCTAACTGACGCCACG	2220
Qy	2221	AGCGGCAAGCAGCTAGACCTGTCTTCTCGGTGACCTTACCCCAAGGACTGTGGAATTG	2280
Db	2221	AGCGGCAAGCAGCTAGACCTGTCTTCTCGGTGACCTTACCCCAAGGACTGTGGAATTG	2280
Qy	2281	ACTGTCACTCTCATCTGCAAGCGGTGGAGAGTGAAGTCTTACTGTGTCTGTGCGGCTC	2340
Db	2281	ACTGTCACTCTCATCTGCAAGCGGTGGAGAGTGAAGTCTTACTGTGTCTGTGCGGCTC	2340
Qy	2341	ATCATTTGTGTGTGAAAAGAGAAAAGAAACAAACAAAGGGCCCCGTGTGGGTATC	2400
Db	2341	ATCATTTGTGTGTGAAAAGAGAAAAGAAACAAACAAAGGGCCCCGTGTGGGTATC	2400
Qy	2401	TACATGGCAACATCAATATCTGAGATGCGAGGCGAGCCAAAAGTTTCAAGAAAGGCGA	2460
Db	2401	TACATGGCAACATCAATATCTGAGATGCGAGGCGAGCCAAAAGTTTCAAGAAAGGCGA	2460
Qy	2461	AAGGCAATGACTCCCATGTGTATGCACTATCGAGACACCATGTGTATATGGCACTGTG	2520
Db	2461	AAGGCAATGACTCCCATGTGTATGCACTATCGAGACACCATGTGTATATGGCACTGTG	2520
Qy	2521	CTACAGGATTCAGAGGGCTCTTCTCTGTAGCGAGAGGTGACACTTACCGGCGTCTCAG	2580
Db	2521	CTACAGGATTCAGAGGGCTCTTCTCTGTAGCGAGAGGTGACACTTACCGGCGTCTCAG	2580
Qy	2581	GGCAACATGGGGGTCTGTCTCTCCCTCCCAACCAACATATCTCAGGGCCCCCAACTGCA	2640
Db	2581	GGCAACATGGGGGTCTGTCTCTCCCTCCCAACCAACATATCTCAGGGCCCCCAACTGCA	2640
Qy	2641	AAGTTGGCATTGAGAGAGCACTTCTGTCTCCCTCTGTAGTGTGAGATGAACCTGATC	2700
Db	2641	AAGTTGGCATTGAGAGAGCACTTCTGTCTCCCTCTGTAGTGTGAGATGAACCTGATC	2700

QY	2701	ACCTTTCCTCCATCCCAACATGGGGATGTAAAGCAGCAGGACACAGCATTCCTCTTA	2760
Db	2701	ACCTTTCCTCCATCCCAACATGGGGATGTAAAGCAGCAGGACACAGCATTCCTCTTA	2760
QY	2761	AAACCTCAGGAGGCCCATGGAGCCGACGAGAAATATCTGAATCCATTCGACGCTTGTG	2820
Db	2761	AAACCTCAGGAGGCCCATGGAGCCGACGAGAAATATCTGAATCCATTCGACGCTTGTG	2820
QY	2821	GTTCCTAAAGCAGGAGCCTGAGACACCCCGTCGTGTCTTAACAGAAATTCCTAAAGAA	2880
Db	2821	GTTCCTAAAGCAGGAGCCTGAGACACCCCGTCGTGTCTTAACAGAAATTCCTAAAGAA	2880
QY	2881	GAGGAATTTATACAGAAAGAAACAGCAGGAGGTTTTCTGGAACACCGCCACTTCACATTC	2940
Db	2881	GAGGAATTTATACAGAAAGAAACAGCAGGAGGTTTTCTGGAACACCGCCACTTCACATTC	2940
QY	2941	TCAGTGAATCTATCTTAAGGGCAAGACATTTGAATATGAATTCGAATCTGAATACGT	3000
Db	2941	TCAGTGAATCTATCTTAAGGGCAAGACATTTGAATATGAATTCGAATCTGAATACGT	3000
QY	3001	CATGACAGCTCATGTGCTCTCACTTAAGGCTGCGGTTAGCAGGCTGTAAATGAGAG	3060
Db	3001	CATGACAGCTCATGTGCTCTCACTTAAGGCTGCGGTTAGCAGGCTGTAAATGAGAG	3060
QY	3061	AGAGAGCCTTAGTACACTAGCATAAGGGTTGACAGCAAGCCCTGATTCAGAGTTAAAC	3120
Db	3061	AGAGAGCCTTAGTACACTAGCATAAGGGTTGACAGCAAGCCCTGATTCAGAGTTAAAC	3120
QY	3121	AGAGGCTTGCCCTCTTCAGAGCAACAGTTCCAAATTCGAAGAGCTACCTGAGTCCCTA	3180
Db	3121	AGAGGCTTGCCCTCTTCAGAGCAACAGTTCCAAATTCGAAGAGCTACCTGAGTCCCTA	3180
QY	3181	CTCTCACTGGGGTCCCCAGATGAAGAAAGCAAAATGCTCTTTTATTAATTAATTTATGG	3240
Db	3181	CTCTCACTGGGGTCCCCAGATGAAGAAAGCAAAATGCTCTTTTATTAATTAATTTATGG	3240
QY	3241	TGGTCTGTGTATTTAAAGATCAAAATGTATTAACAACCTAGCTCTTTTCACTGACTTA	3300
Db	3241	TGGTCTGTGTATTTAAAGATCAAAATGTATTAACAACCTAGCTCTTTTCACTGACTTA	3300
QY	3301	GTAATTAATCTATCTAACTGTTTGGATGCTCGGGTTGTGACTTCTACGACGCTAGAT	3360
Db	3301	GTAATTAATCTATCTAACTGTTTGGATGCTCGGGTTGTGACTTCTACGACGCTAGAT	3360
QY	3361	AAAGGTGTCCTGCCCCCAGGTGGTGGAAATTTAAACAATCTGTCCAAACGAAAGAA	3420
Db	3361	AAAGGTGTCCTGCCCCCAGGTGGTGGAAATTTAAACAATCTGTCCAAACGAAAGAA	3420
QY	3421	ATGTGTGTGTTGAGCAGCAATGACATATCTGCTTTGAATGAAGAGACTTCCTGATCTC	3480
Db	3421	ATGTGTGTGTTGAGCAGCAATGACATATCTGCTTTGAATGAAGAGACTTCCTGATCTC	3480
QY	3481	TAGGTGCGTTCTGTGGTATATCCATTTGTGAAATCATCTTGAATCCATTTGCTCTAATGT	3540
Db	3481	TAGGTGCGTTCTGTGGTATATCCATTTGTGAAATCATCTTGAATCCATTTGCTCTAATGT	3540
QY	3541	CTTGAAGCAATAGAAATTTCTCAATTTCCATGTCGGGTTCTCTAGCTGACGACATA	3600
Db	3541	CTTGAAGCAATAGAAATTTCTCAATTTCCATGTCGGGTTCTCTAGCTGACGACATA	3600
QY	3601	CTTTGACATTTAAAGAGAAATTTGAGAAATATCTCATCTCTTAATAATGTTTAAATATA	3660
Db	3601	CTTTGACATTTAAAGAGAAATTTGAGAAATATCTCATCTCTTAATAATGTTTAAATATA	3660
QY	3661	TACCAAAACAGTGCCCTCGATTAAGTTTCTGTGGCACTGCAACCCATTACTGTGTAG	3720
Db	3661	TACCAAAACAGTGCCCTCGATTAAGTTTCTGTGGCACTGCAACCCATTACTGTGTAG	3720
QY	3721	CTTAAAAACAACAATTAAGCTTAATAGTCCTGGGATCAGAAATTCGAAATGAGATGTCCT	3780
Db	3721	CTTAAAAACAACAATTAAGCTTAATAGTCCTGGGATCAGAAATTCGAAATGAGATGTCCT	3780

Qy	3781	GAATGAAAATCAAGGTCAGAGAGCTGTGCTCTCTTGAAAGGCTATAGGGAAGAAAGCG	3840
Db	3781	GAATGAAAATCAAGGTCAGAGAGCTGTGCTCTCTTGAAAGGCTATAGGGAAGAAAGCG	3840
Qy	3841	GTTCCCTTGGCAATTTCAAGCTTTTAAGAGCGTGAGCTCATTTCCAGAGCTCAGTGGCTGGTC	3900
Db	3841	GTTCCCTTGGCAATTTCAAGCTTTTAAGAGCGTGAGCTCATTTCCAGAGCTCAGTGGCTGGTC	3900
Qy	3901	AAGCTTTTCTCAATGCGATCATCTGTGACATGCGCCCTTCCACTTCCCTTTGACTTAC	3960
Db	3901	AAGCTTTTCTCAATGCGAGCTCATCTGTGACATGCGCCCTTCCACTTCCCTTTGACTTAC	3960
Qy	3961	AAAGCCACACAGAGAGATCCAGAGATTAATCTCCATCTAAAGATCCTTCATCATCTGGA	4020
Db	3961	AAAGCCACACAGAGAGATCCAGAGATTAATCTCCATCTAAAGATCCTTCATCATCTGGA	4020
Qy	4021	AGAGCTTTTGGCATGCAAGACAATATAGCCACAGTGGGGATTATAGACAGACATCTT	4080
Db	4021	AGAGCTTTTGGCATGCAAGACAATATAGCCACAGTGGGGATTATAGACAGACATCTT	4080
Qy	4081	TGGGGTGCTGTTATTCTGCTTACACACCTTCTGSCCATGACTTCCACAGAGAGGCTA	4140
Db	4081	TGGGGTGCTGTTATTCTGCTTACACACCTTCTGSCCATGACTTCCACAGAGAGGCTA	4140
Qy	4141	CAAAATGATCTGGCCGACAGGGATGTTTTGTTTACTGTCGGAGCTTAAACTTTAAAAA	4200
Db	4141	CAAAATGATCTGGCCGACAGGGATGTTTTGTTTACTGTCGGAGCTTAAACTTTAAAAA	4200
Qy	4201	ACCCAGATCAGAAAGATCTGGGCATGCTGGGGCTCACAATTCATCCTTGAACAACATGGC	4260
Db	4201	ACCCAGATCAGAAAGATCTGGGCATGCTGGGGCTCACAATTCATCCTTGAACAACATGGC	4260
Qy	4261	TGAGCTGGGCAACAAGCTCTGCTTTAGAGAGGGATGTCACTTCAACAGTTCACACAGC	4320
Db	4261	TGAGCTGGGCAACAAGCTCTGCTTTAGAGAGGGATGTCACTTCAACAGTTCACACAGC	4320
Qy	4321	CCACACTAGCCCTTATCATTTCCCAATGAGGCTTAAGTGTGTTTCTTACTGATCAATG	4380
Db	4321	CCACACTAGCCCTTATCATTTCCCAATGAGGCTTAAGTGTGTTTCTTACTGATCAATG	4380
Qy	4381	CCCTGACAGTTGCAATTTATGTAATGAAAAAGAAAGACTGGGATTAATCTCTATCAGG	4440
Db	4381	CCCTGACAGTTGCAATTTATGTAATGAAAAAGAAAGACTGGGATTAATCTCTATCAGG	4440
Qy	4441	TGAGTAGACATGAGACCAATGTGTGCTCACAATTACCTTTTCTTTTCTTTTCT	4500
Db	4441	TGAGTAGACATGAGACCAATGTGTGCTCACAATTACCTTTTCTTTTCTTTTCT	4500
Qy	4501	TTTTCTTTTTTTTTTAAATGTGAGACAGGATCTCATTTCTGTTGCTTAGGCTGAGTGCAG	4560
Db	4501	TTTTCTTTTTTTTTTAAATGTGAGACAGGATCTCATTTCTGTTGCTTAGGCTGAGTGCAG	4560
Qy	4561	TGGCGCAATCTGGGCTCATGTGAACCTGTGCTCTGGGGCTCAAGCATTCGCCACCTC	4620
Db	4561	TGGCGCAATCTGGGCTCATGTGAACCTGTGCTCTGGGGCTCAAGCATTCGCCACCTC	4620
Qy	4621	AGCTCTCCAAATAGCTGGGATCACTGGGCAACAACACATGCGCCAGCTTAATTTGTATTT	4680
Db	4621	AGCTCTCCAAATAGCTGGGATCACTGGGCAACAACACATGCGCCAGCTTAATTTGTATTT	4680
Qy	4681	TTTGTAGAGACAGGGTTTCAACATGTGGCCAGGCTGTCTTCAACTCTGCGGCTCAAGC	4740
Db	4681	TTTGTAGAGACAGGGTTTCAACATGTGGCCAGGCTGTCTTCAACTCTGCGGCTCAAGC	4740
Qy	4741	AATCTCTCTGCTCGGCTTCCAAAGTGTGGGATTTACAGATGTGAGCAACCGATCTCAG	4800
Db	4741	AATCTCTCTGCTCGGCTTCCAAAGTGTGGGATTTACAGATGTGAGCAACCGATCTCAG	4800
Qy	4801	CCCCACACCTCATTTATACAAATTAACGCGCCAGTAACTGTGACTTTTGCTTCTCAC	4860
Db	4801	CCCCACACCTCATTTATACAAATTAACGCGCCAGTAACTGTGACTTTTGCTTCTCAC	4860
Qy	4861	CCCTGCTGATCTGGAGAGAGGGAATTATGTTAAGCTTGTCAACAAGTCCCAAGTT	4920

Db	4861	CCCTGCTCTGATCTGGAAAGAGAGGATTAATGTTATAGCTTGTACGACAGTCCCAAGTT	4920
Qy	4921	CAATATTTCTGCGGCAAAACTTCTTCAAAAAATAATGTAATCTACTTCTATTTCAATGA	4980
Db	4921	CAATATTTCTGCGGCAAAACTTCTTCAAAAAATAATGTAATCTACTTCTATTTCAATGA	4980
Qy	4991	ATTCACTCTGGAAATGACCGGCTCCACTTGTTCACATGGCACTAAATGAAGAATTTTA	5040
Db	4991	ATTCACTCTGGAAATGACCGGCTCCACTTGTTCACATGGCACTAAATGAAGAATTTTA	5040
Qy	5041	TAGTCTCTTAATGSGCTGTACTGCAAGACCTCTYGAACCTTCCAGAGATAGATAT	5100
Db	5041	TAGTCTCTTAATGSGCTGTACTGCAAGACCTCTYGAACCTTCCAGAGATAGATAT	5100
Qy	5101	TTAAGTCAATGCGCTTGCGGCTTGCTATGGACCTTTCCTTGTGAAAGTCTGTCTCTGC	5160
Db	5101	TTAAGTCAATGCGCTTGCGGCTTGCTATGGACCTTTCCTTGTGAAAGTCTGTGTCTGC	5160
Qy	5161	CCAAGTACCTTTGGGCTTTGTGAGCCGAGATGCTGACCTCTGATTAAGGCGCAAGAGGG	5220
Db	5161	CCAAGTACCTCTTGCGCTTGTGAGCCGAGATGCTGACCTCTGATTAAGGCGCAAGAGGG	5220
Qy	5221	CTGGGCGCTTCCTTCCCTACATGAAAGACCCCTTAATTGAATCACTGTGTGAGCCTAGC	5280
Db	5221	CTGGGCGCTTCCTTCCCTACATGAAAGACCCCTTAATTGAATCACTGTGTGAGCCCTAGC	5280
Qy	5281	CCCTCATTTCTGCATTTCCCAACCTTCCAGACCCCTTCCAAAGAGACTAGTGCCCTGC	5340
Db	5281	CCCTCATTTCTGCATTTCCCAACCTTCCAGACCCCTTCCAAAGAGAGACTAGTGCCCTGC	5340
Qy	5341	ATTCCACCCCAAGGTGGGATTTGGCTTCTTAAGCTGTGCTACTGTCACTACCGACAT	5400
Db	5341	ATTCCACCCCAAGGTGGGATTTGGCTTCTTAAGCTGTGCTACTGTCACTACCGACAT	5400
Qy	5401	CACGTGGCTCTGCAAGGACACACGTGGCCATTTCTTAACTGAGGGGCTCAAAATCTCC	5460
Db	5401	CACGTGGCTCTGCAAGGACACACGTGGCCATTTCTTAACTGAGGGGCTCAAAATCTCC	5460
Qy	5461	TGGAACAAGTTCGTGGCTCTGAGACACAGTATTTTCTGGAAGCTGTGCTCAGTGAAGGGG	5520
Db	5461	TGGAACAAGTTCGTGGCTCTGAGACACAGTATTTTCTGGAAGCTGTGCTCAGTGAAGGGG	5520
Qy	5521	CCAGCTTGAGGAACCTTGCGCTCTTTTCTTTAAAGCCAGGCCCCCACTTAACATAAATTT	5580
Db	5521	CCAGCTTGAGGAACCTTGCGCTCTTTTCTTTAAAGCCAGGCCCCCACTTAACATAAATTT	5580
Qy	5581	TCAGGGTCACTGGAAACAGTGAAGTGCATTTGTGAAGCTTACTGATCCAGCCCACT	5640
Db	5581	TCAGGGTCACTGGAAACAGTGAAGTGCATTTGTGAAGCTTACTGATCCAGCCCACT	5640
Qy	5641	GCTCATCAAGTGTCTGCATGCTTACGAGAGGCGACGCGCATGAGAGACTGTGCTCT	5700
Db	5641	GCTCATCAAGTGTCTGCATGCTTACGAGAGGCGACGCGCATGAGAGACTGTGCTCT	5700
Qy	5701	AATGCTGTGTCATTTGACAGAAAGGAAGGTCTCAAGAAAGTCAACTGGGCAAGCA	5760
Db	5701	AATGCTGTGTCATTTGACAGAAAGGAAGGTCTCAAGAAAGTCAACTGGGCAAGCA	5760
Qy	5761	CAAGCCCAACCGGACATGGCTTGTGTAAGGTGACACACTGTGTGTGTGATCTGCACT	5820
Db	5761	CAAGCCCAACCGGACATGGCTTGTGTAAGGTGACACACTGTGTGTGTGATCTGCACT	5820
Qy	5821	GCTTCACGTGAAATATTTATTCATTTGACAGATCTTTTAAAGTGGCACTTTATTCATTT	5880
Db	5821	GCTTCACGTGAAATATTTATTCATTTGACAGATCTTTTAAAGTGGCACTTTATTCATTT	5880
Qy	5881	CTGTGCTTTAAATTAACAATGTACCAAAAAACAAGTATCAAGCTGTTTAAGTCTTCGG	5940
Db	5881	CTGTGCTTTAAATTAACAATGTACCAAAAAACAAGTATCAAGCTGTTTAAGTCTTCGG	5940
Qy	5941	CTACTGTGCTGCTGTTCAGTAGAGGCCCGGTTTCCAGTGTGTGACTGTGACAGGCTC	6000

Db	5941	CTACTGTGCCCTGGTTCAGTAGAGAGCCCGGTTCCAGATTGTTGACTGACAGGCTC	6000
Qy	6001	AGCATGGGCTCAGCAGATGCTGCTCTTAATTTGTGATGATATCAGAAAGCCAGGCTTTGGG	6060
Db	6001	AGCATGGGCTCAGCAGATGCTGCTCTTAATTTGTGATGATATCAGAAAGCCAGGCTTTGGG	6060
Qy	6061	ATACAAAGTCTTTCCTCTTCATTTGATGATCGGCACTGCTGGAAGCAGATGTTTTCGTC	6120
Db	6061	ATACAAAGTCTTTCCTCTTCATTTGATGATCGGCACTGCTGGAAGCAGATGTTTTCGTC	6120
Qy	6121	GGAAATATAAAATAATAGTCTTGGAGTCTGGCCAAAAAATAAAAAA	6163
Db	6121	GGAAATATAAAATAATAGTCTTGGAGTCTGGCCAAAAAATAAAAAA	6163
RESULT 2			
ABA99506			
ID	ABA99506	standard; cDNA; 5897 BP.	
AC	ABA99506;		
XX	17-MAY-2002	(first entry)	
DE	Human tumour-associated antigen B345 cDNA.		
KW	Tumour-associated antigen; human; B345; cytostatic; cell communication;		
KW	cell interaction; signal transduction; metastasis; cancer; colon;		
KW	immunotherapy; carcinoma; lung; diagnosis; gene; ss.		
OS	Homo sapiens.		
XX			
FT	Key	Location/Qualifiers	
FT	5'UTR	1..214	
FT		/*tag= a	
FT	CDS	215..2464	
FT		/*tag= b	
FT		/product= "tumour-associated antigen B345"	
FT	3'UTR	2465..5897	
FT		/*tag= c	
PN	WO200204508-A1.		
XX			
PD	17-JAN-2002.		
XX			
PP	05-JUL-2001; 2001WO-EP07705.		
XX			
PR	07-JUL-2000; 2000DE-1033080.		
PR	19-APR-2001; 2001DE-1019294.		
XX	(BOEH ) BOEHRINGER INGELHEIM INT GMBH.		
XX			
PI	Schweifer N, Scherl-Mostageer M, Sommergruber W, Abseher R;		
DR	WPI: 2002-171704/22.		
XX	P-PSDB; NAM49640.		
XX			
XX	New tumor-associated antigen B345, useful for diagnosis and		
PT	immunotherapy of tumors, also related nucleic acid and antibodies		
PS	Example 6; Page 70-76; 102pp; German.		
XX			
CC	This invention describes a novel tumour-associated antigen, designated		
CC	B345 which has cytostatic activity. B345 is involved in communication,		
CC	intersection and/or signal transduction with extracellular components and		
CC	ligands, especially in the metastatic potential of cancers, particularly		
CC	of the colon. B345 or its immunogenic fragments, also the DNA that		
CC	encodes it, are useful for immunotherapy of cancer, particularly		
CC	carcinoma of lung or colon. Antibodies raised against B345 are useful for		
CC	treatment and diagnosis of cancers that are associated with B345		
CC	expression, including their use for targeted delivery of cytotoxic or		
CC	radioactive agents. Probes derived from B345 can be used to detect		
CC	tumour-specific mutations in the B345 sequence, and can be used to screen		
CC	for B345 specific modulators. This sequence encodes a human B345		

CC	tumour-associated antigen described in the invention.
XX	
SQ	Sequence 5897 BP; 1464 A; 1596 C; 1364 G; 1467 T; 6 other;
	87.5%; Score 5395; DB 24; Length 5897;
	Query Match Best Local Similarity 99.9%; Pred. No. 0;
	Matches 5865; Conservative 0; Mismatches 7; Indels 1; Gaps 1;
OY	291 CCTGAACCTGCGGGGTCTCTATCGCATGCTAGAGGGTTCCTGCTGGGTGCGGCGCCT 350
Dd	25 CCTGAACCTGCGGGGTCTCTATCGCATGCTAGAGGGTTCCTGCTGGGTGCGGCGCCT 84
OY	351 GCCCGCGGGGCAGAAGCTTTTGAGATTGCTCTGCCACGAAAGCAACATTACAGTTCT 410
Dd	85 GCCCGCGGGGCAGAAGCTTTTGAGATTGCTCTGCCACGAAAGCAACATTACAGTTCT 144
OY	411 CATTAAGCTGGGGACCCCGACTGCTGGCAAAACCCTGTACATCGTCATTCTTAAG 470
Dd	145 CATTAAGCTGGGGACCCCGACTGCTGGCAAAACCCTGTACATCGTCATTCTTAAG 204
OY	471 ACATATAACATGTTGTCATCAAAGTCTGGAAAGAAATAGCTTTTACCTTAGCTGCA 530
Dd	205 ACATATAACATGTTGTCATCAAAGTCTGGAAAGAAATAGCTTTTACCTTAGCTGCA 264
OY	531 GAGTCCGAGAAATCACTTTGTCTATAGAGATCCAAAAAATTTAGACTATAGTCAAGCCC 590
Dd	265 GAGTCCGAGAAATCACTTTGTCTATAGAGATCCAAAAAATTTAGACTATAGTCAAGCCC 324
OY	591 ATGTCCTTTTTGGGAGGTTGACGCTTCAAGCCCTCGACAATCGTTGTGCTTACCTCAACAG 650
Dd	325 ATGTCCTTTTTGGGAGGTTGACGCTTCAAGCCCTCGACAATCGTTGTGCTTACCTCAACAG 384
OY	651 AACTTTCATCTGGAGTGTCAAAGCTCATTAAGAGCATCGGTTTAGAGCTGACGATTTTCAT 710
Dd	385 AACTTTCATCTGGAGTGTCAAAGCTCATTAAGAGCATCGGTTTAGAGCTGACGATTTTCAT 444
OY	711 CCCTTCGCTGAGGCAATCGGTCCGGGTGAGAGCTGCCAAGCGAGTCACTCACTCAT 770
Dd	445 CCCTTCGCTGAGGCAATCGGTCCGGGTGAGAGCTGCCAAGCGAGTCACTCACTCAT 504
OY	771 CAGGAGCGCAATGCATGCAACCGTGTGAGGATCGGAAACCTTCTGACGCAATGCACTGT 830
Dd	505 CAGGAGCGCAATGCATGCAACCGTGTGAGGATCGGAAACCTTCTGACGCAATGCACTGT 564
OY	831 GTCCCGGATCAAGATGCAAGAGAGAGTGAATAATGGCTTACACCTCCATGGTTCACCC 890
Dd	565 GTCCCGGATCAAGATGCAAGAGAGAGTGAATAATGGCTTACACCTCCATGGTTCACCC 624
OY	891 CAGAAATGTCTCCGGCTTCAGCATTTGCAAAACCGCTCATCTATAAAGTCTGTGCATCAT 950
Dd	625 CAGAAATGTCTCCGGCTTCAGCATTTGCAAAACCGCTCATCTATAAAGTCTGTGCATCAT 684
OY	951 CGAATCTGTGTTTAAAGGTTGAAGGCTCAGCAACCTGATGTCTGCCAACTAACCGAAGG 1010
Dd	685 CGAATCTGTGTTTAAAGGTTGAAGGCTCAGCAACCTGATGTCTGCCAACTAACCGAAGG 744
OY	1011 CTTCCTGAGAGATGAGCTCATGAGTGGAGTGTGTGCTTCTGCAACCTGCGGGCGAG 1070
Dd	745 CTTCCTGAGAGATGAGCTCATGAGTGGAGTGTGTGCTTCTGCAACCTGCGGGCGAG 804
OY	1071 CGTCTCTCTTCTCAACTTCAACCTCTTCCAACCTGTGAGAGGAAAGGAGCGGGTGGATA 1130
Dd	805 CGTCTCTCTTCTCAACTTCAACCTCTTCCAACCTGTGAGAGGAAAGGAGCGGGTGGATA 864
OY	1131 CTAATATCCCGGGCTTCCACACCAACCCCGAGGTGTTCAGCTGGAAGCAACAGACCTGG 1190
Dd	865 CTAATATCCCGGGCTTCCACACCAACCCCGAGGTGTTCAGCTGGAAGCAACAGACCTGG 924
OY	1191 GAACATGGCGGGGAACCTTCAACCTCTCTCTGCAAGGCTGTGACCAAGATGCCAAAATGCC 1250
Dd	925 GAACATGGCGGGGAACCTTCAACCTCTCTCTGCAAGGCTGTGACCAAGATGCCAAAATGCC 984
OY	1251 AGGATCCTCCGGCTGCAATTCCAAGTTTGGTTCCAATTCACMAAATGAAGCAATTA 1310

Db 985 AGGATCCCTCCGCGCTGACGATTCGCAAGTTTGGTCCAACTCCACAAAGAAAGCAATTA 1044  
Qy 1311 AATCAACGTGGTGTGACTTGATGATGAGGAGCATGTCACTACCATCGAGCCAGGCC 1370  
Db 1045 AATCAACGTGGTGTGACTTGATGATGAGGAGCATGTCACTACCATCGAGCCAGGCC 1104  
Qy 1371 CGTCAACAGAGCCGCAAGTTTGTCTGGCTGTGTCTGTGTCTGAATCTCGAGCTG 1430  
Db 1105 CGTCAACAGAGCCGCAAGTTTGTCTGGCTGTGTCTGTGTCTGAATCTCGAGCTG 1164  
Qy 1431 CAGTGAACCTCACCCCTGACATCTGGCTCCAAACGAAATCTCCCTTCCTTTTGAATGA 1490  
Db 1165 CAGTGAACCTCACCCCTGACATCTGGCTCCAAACGAAATCTCCCTTCCTTTTGAATGA 1224  
Qy 1491 TGTGACACGTCTGTGGATGAATGTGAAAAAACAATTAAGCTGACAGACACCGGTAAG 1550  
Db 1225 TGTGACACGTCTGTGGATGAATGTGAAAAAACAATTAAGCTGACAGACACCGGTAAG 1284  
Qy 1551 CCAAGAGAAATCTCACTCACTCCAGGTGCGCAGTGCATCTCCACCTGCTGTGGAGCT 1610  
Db 1285 CCAAGAGAAATCTCACTCACTCCAGGTGCGCAGTGCATCTCCACCTGCTGTGGAGCT 1344  
Qy 1611 GATGACTTCTCTGGAAGCTGTGTGCTGCTCCAAAGACAGGCTCAGCTGTGTGTGTGCTC 1670  
Db 1345 GATGACTTCTCTGGAAGCTGTGTGCTGCTCCAAAGACAGGCTCAGCTGTGTGTGTGCTC 1404  
Qy 1671 AGCCCGAAGCTGACAGCAGCATACACAGAGAAAGCCCTGCAACACAGGCTTACGTAAGCT 1730  
Db 1405 AGCCCGAAGCTGACAGCAGCATACACAGAGAAAGCCCTGCAACACAGGCTTACGTAAGCT 1464  
Qy 1731 CGTGGCAGTGCATACCCAGCCAGGACCTGTATCTTGCGCTCTTCTGCGCGGAGGCTC 1790  
Db 1465 CGTGGCAGTGCATACCCAGCCAGGACCTGTATCTTGCGCTCTTCTGCGCGGAGGCTC 1524  
Qy 1791 TATCAAGCAGATCCAGGTGAAGCAAAATCTCGGTGACCCCTTCGACCTTTGCCCCAG 1850  
Db 1525 TATCAAGCAGATCCAGGTGAAGCAAAATCTCGGTGACCCCTTCGACCTTTGCCCCAG 1584  
Qy 1851 CTTCCCAACAAAGAGGCTCTCAGGACGGGTCTGAAGGCTCTTATTAATCACTTATTTCAAA 1910  
Db 1585 CTTCCCAACAAAGAGGCTCTCAGGACGGGTCTGAAGGCTCTTATTAATCACTTATTTCAAA 1644  
Qy 1911 GGAAGGCGTTTTCACGGTGACCCCTGACACAAAGAAAGGCTTACCTGAGACCCCCAA 1970  
Db 1645 GGAAGGCGTTTTCACGGTGACCCCTGACACAAAGAAAGGCTTACCTGAGACCCCCAA 1704  
Qy 1971 CTGGGACCCGGGGCTGCTCCATCTCTGCTGTGTCTGGAACATAGGCTGCCAGAA 2030  
Db 1705 CTGGGACCCGGGGCTGCTCCATCTCTGCTGTGTCTGGAACATAGGCTGCCAGAA 1764  
Qy 2031 CCAGGTGGGCTGCTGACTTTCTTTAAGAGCGGAGCGGTGCTGTCAGACAGGGCG 2090  
Db 1765 CCAGGTGGGCTGCTGACTTTCTTTAAGAGCGGAGCGGTGCTGTCAGACAGGGCG 1824  
Qy 2091 CGCATTCATGATCATCCAGAGACGCGAACCCGGGCTGAGAGATCTTACGCTGAGCGA 2150  
Db 1825 CGCATTCATGATCATCCAGAGACGCGAACCCGGGCTGAGAGATCTTACGCTGAGCGA 1884  
Qy 2151 GGATGTGCTCCCAAGACGCTTCCACATCAAGACTTCTGGGTCAACATCTTAACG 1944  
Db 1885 GGATGTGCTCCCAAGACGCTTCCACATCAAGACTTCTGGGTCAACATCTTAACG 1944  
Qy 2211 CAGCCCCAGAGGGGAGAGAGCTAGACCGGCTCTTCTGGGTGACATTAACCCAAAGAG 2270  
Db 1945 CAGCCCCAGAGGGGAGAGAGCTAGACCGGCTCTTCTGGGTGACATTAACCCAAAGAG 2004  
Qy 2271 TGTGACTTGAATGATCTCATCTGACAGCGGTGAGAGGTGAAGTCTTAAGTGTCTGCTC 2330  
Db 2005 TGTGACTTGAATGATCTCATCTGACAGCGGTGAGAGGTGAAGTCTTAAGTGTCTGCTC 2064  
Qy 2331 CCTGGGCTCATCTTGTCTGTGTGAAAAAGAAAGAAAGACAAACAGGGCCCGC 2390

Db 2065 CCTGGGCTCATCTTGTCTGTGTGAAAAAGAAAGAAAGACAAACAGGGCCCGC 2124  
Qy 2391 TGTGGATATCTAACAATGCAACATCAATATCTAGATGCGAGGACCCAAAAAGTTTCA 2450  
Db 2125 TGTGGATATCTAACAATGCAACATCAATATCTAGATGCGC-AGGCAAGCCAAAAAGTTTCA 2183  
Qy 2451 GAAAGGGGAAAGGACATGACTCCCATGTGTATATGACTCATTCGAGGACACATGTTATA 2510  
Db 2184 GAAAGGGGAAAGGACATGACTCCCATGTGTATATGACTCATTCGAGGACACATGTTATA 2243  
Qy 2511 TGGGATCTGCTACAGATATTCAGGCGGCTCTTCTGACGCGAGGTGACACCTTACCG 2570  
Db 2244 TGGGATCTGCTACAGATATTCAGGCGGCTCTTCTGACGCGAGGTGACACCTTACCG 2303  
Qy 2571 GCCGTTCCAGGGCACATGCGGGGTGTCTCTCCCTCCCAACCATATGCTCCAGGCGC 2630  
Db 2304 GCCGTTCCAGGGCACATGCGGGGTGTCTCTCCCTCCCAACCATATGCTCCAGGCGC 2363  
Qy 2631 CCCAATCGAAAGTTGGCCATCGAGGAGCCACCTCTGCTCCCTCTGAGTCTGAGAG 2690  
Db 2364 CCCAATCGAAAGTTGGCCATCGAGGAGCCACCTCTGCTCCCTCTGAGTCTGAGAG 2423  
Qy 2691 TGAACCGTACACCTTCTCCCATCCCAACAATGGGGATGTAAAGCAGACAGACAT 2750  
Db 2424 TGAACCGTACACCTTCTCCCATCCCAACAATGGGGATGTAAAGCAGACAGACAT 2483  
Qy 2751 TCCCTTACTGAACTCTCAGAGCCCATGGAGCCAGACAGATTAATCTTATCTCAGAC 2810  
Db 2484 TCCCTTACTGAACTCTCAGAGCCCATGGAGCCCATGGAGCCAGACAGATTAATCTTATCTCAGAC 2543  
Qy 2811 GCTTTGCTGAGTTTATTAAGAGGCGACCTGACACACCCGTCCGTGTTCTTAACCAAAA 2870  
Db 2544 GCTTTGCTGAGTTTATTAAGAGGCGACCTGACACACCCGTCCGTGTTCTTAACCAAAA 2603  
Qy 2871 TCCTAAGAGAGGAATTAACAGAAAGAACAGCAGAGGTTTCTGTCGACACCGCAAC 2930  
Db 2604 TCCTAAGAGAGGAATTAACAGAAAGAACAGCAGAGGTTTCTGTCGACACCGCAAC 2663  
Qy 2931 TTCACATTGCTCAGTGAACCTCATTTCTAAGGGGCAAGACATTTGAATTTCCAAATC 2990  
Db 2664 TTCACATTGCTCAGTGAACCTCATTTCTAAGGGGCAAGACATTTGAATTTCCAAATC 2723  
Qy 2991 TGGATACAGTCAATGACGCTCATGTGCTCTCACTTAAGGCTGTGCGGTAGCCAGCTC 3050  
Db 2724 TGGATACAGTCAATGACGCTCATGTGCTCTCACTTAAGGCTGTGCGGTAGCCAGCTC 2783  
Qy 3051 TAAAGAGAGAGAGGCGCTGACCTAAGATTAAGGGTGAAGGACCGCTGAAATTCAG 3110  
Db 2784 TAAAGAGAGAGAGGCGCTGACCTAAGATTAAGGGTGAAGGACCGCTGAAATTCAG 2843  
Qy 3111 AGTGTAAACAGAGGCTTGCCCTCTTCAAGACAAAGTTCCAAATTCAGAGGCTTACCT 3170  
Db 2844 AGTGTAAACAGAGGCTTGCCCTCTTCAAGACAAAGTTCCAAATTCAGAGGCTTACCT 2903  
Qy 3171 GAGGTCCCTACTCTCACTGCGGCTCCCAAGATGAAGAGACATGTGCTTTTATTAAT 3230  
Db 2904 GAGGTCCCTACTCTCACTGCGGCTCCCAAGATGAAGAGACATGTGCTTTTATTAAT 2963  
Qy 3231 ATTATTTGGGTGCTGCTGTATTTAAGAGATCAAGATTAACCACTAGCTCTTTC 3290  
Db 2964 ATTATTTGGGTGCTGCTGTATTTAAGAGATCAAGATTAACCACTAGCTCTTTC 3023  
Qy 3291 ACCTGACTGATTAATCACTACTACTGTTTGGATGCTCGGGTGTGACTTCTACTG 3350  
Db 3024 ACCTGACTGATTAATCACTACTACTGTTTGGATGCTCGGGTGTGACTTCTACTG 3083  
Qy 3351 ACCGCTAAGTAAAGTGTGCTGTCTCCCAAGGTGTGGAAATTAATTAATCTGTCCA 3410  
Db 3084 ACCGCTAAGTAAAGTGTGCTGTCTCCCAAGGTGTGGAAATTAATTAATCTGTCCA 3143  
Qy 3411 CCAGAAAGAAATGTGTGCTTTGAAGCAGATTTGAACAATATGCTTTGAATGAAGACCT 3470  
Db 3144 CCAGAAAGAAATGTGTGCTTTGAAGCAGATTTGAACAATATGCTTTGAATGAAGACCT 3203

3471 CCTGATCTCTAGAGTCGGTTCGGTTATCCCATTTGGAAATTCATCTTGAATCCATT 3530  
3204 CCGATTTCTCTAGAGTCGGTTCGGTTATCCCATTTGGAAATTCATCTTGAATCCATT 3263  
3531 GTCTTAATAGTCCTAGCAATTAAGAAATTTCTCAAGTTTCATGTGCGGTTCTCTAGC 3590  
3264 GTCTTAATAGTCCTAGCAATTAAGAAATTTCTCAAGTTTCATGTGCGGTTCTCTAGC 3323  
3591 TGCACCAATCTTGAATTAAGAAATTTAGAAATTTTCATTCCTTAAGAAATG 3650  
3324 TGCACCAATCTTGAATTAAGAAATTTAGAAATTTTCATTCCTTAAGAAATG 3383  
3651 TTTAATATATACCAAAAGTGGCCCTGCTAGTATGTTTCTGTTGCCATGCAACCAT 3710  
3384 TTTAATATATACCAAAAGTGGCCCTGCTAGTATGTTTCTGTTGCCATGCAACCAT 3443  
3711 TACTTGTAGCTTTAAACCAACATTAAGTCTTATAGTCTTGGGGATCAGAAATTCGAAAT 3770  
3444 TACTTGTAGCTTTAAACCAACATTAAGTCTTATAGTCTTGGGGATCAGAAATTCGAAAT 3503  
3771 GGATGTCCTGTAATGAATTAAGGTCAGAGAGTGTGCTCTCTTCTGAAGGCTTAG 3830  
3504 GGATGTCCTGTAATGAATTAAGGTCAGAGAGTGTGCTCTCTTGAAGGCTTAG 3563  
3831 GGAGAAGCCGGTTCCTGTCATTTCAAGCTTCTAGAGGCTGCTGATTCGAGGCTCA 3890  
3564 GGAGAAGCCGGTTCCTGTCATTTCAAGCTTCTAGAGGCTGCTGATTCGAGGCTCA 3623  
3891 GTGGCTGTGAAGCTTTTCTCAATGACATCTGACATCTGGCCCTCCACTTCCTC 3950  
3624 GTGGCTGTGAAGCTTTTCTCAATGACATCTGACATCTGGCCCTCCACTTCCTC 3683  
3951 TTTGACTTAACAAAGCCCAAGAAATTCAGAAATTCCTCCATCTTAAATCTCTTA 4010  
3684 TTTGACTTAACAAAGCCCAAGAAATTCAGAAATTCCTCCATCTTAAATCTCTTA 3743  
4011 TCATCTGGAAGAGCTTTTTCATGCAAGCAACATAGCAAGGTTGGAATTAAGACC 4070  
3744 TCATCTGGAAGAGCTTTTTCATGCAAGCAACATAGCAAGGTTGGAATTAAGACC 3803  
4071 AGGACATCTTTGGGGTGTATTATCTGCTCAACAACCTTCTGCACTGACTCCACA 4130  
3804 AGGACATCTTTGGGGTGTATTATCTGCTCAACAACCTTCTGCACTGACTCCACA 3863  
4131 GGAAGGCTTAACAAATGATCTGCGCAAGGATGTTTGTAGCTTTCGCACTCTAAC 4190  
3864 GGAAGGCTTAACAAATGATCTGCGCAAGGATGTTTGTAGCTTTCGCACTCTAAC 3923  
4191 ACTTAAACCAACCCAGATCAGAAATCTGGCATGCTGGGCTCACTTCACTAGC 4250  
3924 ACTTAAACCAACCCAGATCAGAAATCTGGCATGCTGGGCTCACTTCACTAGC 3983  
4251 AACCACTGGCTGAAGTGGGCAACAGCTTGTGCTTAAAGAGGTTTCACTTCAACAGG 4310  
3984 AACCACTGGCTGAAGTGGGCAACAGCTTGTGCTTAAAGAGGTTTCACTTCAACAGG 4043  
4311 TCACCAAGCCCACTACGCTTATCACTTCCCAATGAGGCTTAAGTGTGTTCTA 4370  
4044 TCACCAAGCCCACTACGCTTATCACTTCCCAATGAGGCTTAAGTGTGTTCTA 4103  
4371 CTGATCAATGCCCCTGCAAGTTCATTTATGTAATGAAGAAAGAACTGGGATTAATC 4430  
4104 CTGATCAATGCCCCTGCAAGTTCATTTATGTAATGAAGAAAGAACTGGGATTAATC 4163  
4431 TCTAATCAGGTAGTACATGAGCAATGATGTGCTCACTTACCTTTTCTTTT 4490  
4164 TCTAATCAGGTAGTACATGAGCAATGATGTGCTCACTTACCTTTTCTTTT 4223  
4491 TTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 4550  
4224 TTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 4283

4551 TGAATGAGTGGGCAATCTCGGCTCACTGCAACTCTGCTCTCGGGCTCAAGCAAT 4610  
4284 TGAATGAGTGGGCAATCTCGGCTCACTGCAACTCTGCTCTCGGGCTCAAGCAAT 4343  
4611 CTCCCACTCAGCTTCCCAATATGTTGGATTCATGGGCAAAACCAATGCTCCAGCTTA 4670  
4344 CTCCCACTCAGCTTCCCAATATGTTGGATTCATGGGCAAAACCAATGCTCCAGCTTA 4403  
4671 TTTTGTATTTTGTAGAGACAGGTTTCACATTTGCTCCAGGCTGGTCAACTCTCT 4730  
4404 TTTTGTATTTTGTAGAGACAGGTTTCACATTTGCTCCAGGCTGGTCAACTCTCT 4463  
4731 GGGCTCAAGCAATCTCTGCTCGGCTCCCAAGTCTGGGATTAAGATGAGGCA 4790  
4464 GGGCTCAAGCAATCTCTGCTCGGCTCCCAAGTCTGGGATTAAGATGAGGCA 4523  
4791 CCGCATCCAGCCCAACCTCTCATTTATTAACAATTAAGTCTGCCAGTAATGAGCTTTT 4850  
4524 CCGCATCCAGCCCAACCTCTCATTTATTAACAATTAAGTCTGCCAGTAATGAGCTTTT 4583  
4851 GCTTCTCAACCCCTGCTGATCTGGAAGAGAGGATTAAGTTATAGCTTGTGAGCA 4910  
4584 GCTTCTCAACCCCTGCTGATCTGGAAGAGAGGATTAAGTTATAGCTTGTGAGCA 4643  
4911 GTCCCAAGTTCAATTTCTGCGGCAAAACTTCTTCAAAAATTAATGTAATTAATG 4970  
4644 GTCCCAAGTTCAATTTCTGCGGCAAAACTTCTTCAAAAATTAATGTAATTAATG 4703  
4971 TATTCATGAATTCACCTTGAATATGCAACCGCTCAACTTGTTCATATGCAATTAATGA 5030  
4704 TATTCATGAATTCACCTTGAATATGCAACCGCTCAACTTGTTCATATGCAATTAATGA 4763  
5031 AGGAATTTTATAGTCTCTTAATAGGCGGTACTGGAACCTTGAACCTTGAACCTTGAAC 5090  
4764 AGGAATTTTATAGTCTCTTAATAGGCGGTACTGGAACCTTGAACCTTGAACCTTGAAC 4823  
5091 GATAGATATTTATAGTCAATGCTTGGCGTTGCTGATGAGCACTTTCCTTCTGAAATG 5150  
4824 GATAGATATTTATAGTCAATGCTTGGCGTTGCTGATGAGCACTTTCCTTCTGAAATG 4883  
5151 TGGTCTGCGCCAGTACCTTGGCTTGTGAGCCGAGATGCTGACCTTGAATTAAGG 5210  
4884 TGGTCTGCGCCAGTACCTTGGCGTTGCTGATGAGCCGAGATGCTGACCTTGAATTAAGG 4943  
5211 CAAGAGAGGCTGCGGCTTCTTCCCTCACTGAAGAGCCCTTATTTGAATTAATGAGG 5270  
4944 CAAGAGAGGCTGCGGCTTCTTCCCTCACTGAAGAGCCCTTATTTGAATTAATGAGG 5003  
5271 GAGCCCTAGCCCTCAATCTGCAACTTCCCAACCTCCAGCCCTTCCAGCAGAGACTA 5330  
5004 GAGCCCTAGCCCTCAATCTGCAACTTCCCAACCTCCAGCCCTTCCAGCAGAGACTA 5063  
5331 GGTGCTCTGCAATTCACCCCAAGTGGATTTGCTTGAAGCTGACTTGTACCA 5390  
5064 GGTGCTCTGCAATTCACCCCAAGTGGATTTGCTTGAAGCTGACTTGTACCA 5123  
5391 TCACCGAATCACTGTTCCTGCAAGGACCAACAGTGGCAATTTCTTCAACTGAAGG 5450  
5124 TCACCGAATCACTGTTCCTGCAAGGACCAACAGTGGCAATTTCTTCAACTGAAGG 5183  
5451 TCACCAATCTGCAAGTGTGCTCTGAGACCAATTTCTTGAAGCTGAGCTCA 5510  
5184 TCACCAATCTGCAAGTGTGCTCTGAGACCAATTTCTTGAAGCTGAGCTCA 5243  
5511 GTGAAGGAGCCCAAGCTGAGAACCTTGTCTTTTCTTAAAGCCAGGCCCACTTAC 5570  
5244 GTGAAGGAGCCCAAGCTGAGAACCTTGTCTTTTCTTAAAGCCAGGCCCACTTAC 5303  
5571 ATTAACATTTCAAGGTACCTGGAACAGTGAAGGCAATTTGTGAAGCTTACGATG 5630  
5304 ATTAACATTTCAAGGTACCTGGAACAGTGAAGGCAATTTGTGAAGCTTACGATG 5363  
5631 CCAGCCACTGCTCATCCAGTGTCTGCTCAATGCTTCAAGAGAGGCGAGGCAATGAG 5690

Db	5364	CCAGCCCACTGCTATCCACGTTGATCCATGCTTACGAGAAAGGCCAGCGCATGACGG	5423
QY	5691	ACTGCTCTTAATGCTGTGTCATTTGCA CAGAAGGAAAAGSTCTTCAGAAAGATCAACT	5750
Db	5424	ANTGGTCTCTAATGNTGTGTTCATTGCA CAGAAAGGAAAGGTCTTCAGAAAGATCAACT	5483
QY	5751	GGGACAAGACACAGGCCCAACGGCA CATGAGCTTTGGTAAAGTTAGACACTGCTGTGTG	5810
Db	5484	GGGACACAGCACAGGCCCAACGGCA CATGAGCTTTGGTAAAGTTAGACACTGCTGTGTG	5543
QY	5811	GATCTGCAGTGCCTTCACTGAGAAATATTTATTCATTCAGATPACTTTTAAAGTGGCAATTT	5870
Db	5544	GATCTGCAGTGCCTTCACTGAGAAATATTTATTCATTCAGATPACTTTTAAAGTGGCAATTT	5603
QY	5871	TATTCATTTCTGTGCTTTAAATTAACCAATGTACCAAAAAACAAGTATCAAGTGTTTA	5930
Db	5604	TATTCATTTCTGTGCTTTAAATTAACCAATGTACCAAAAAACAAGTATCAAGTGTTTA	5663
QY	5931	AGTGCCTGGGCTACTGTGCCCCGTGGTTAGTAGAGAGCCCGGTTTTCCAGTTGTTGACTG	5990
Db	5664	AGTGCCTGGGCTACTGTGCCCCGTGGTTAGTAGAGAGCCCGGTTTTCCAGTTGTTGACTG	5722
QY	5991	TGACAGGCTCAGCATGGGCTCAGCAGATGCTGCTTTAAATTTGTGATGATACAGAAACC	6050
Db	5724	TGACAGGCTCAGCATGGGCTCAGCAGATGCTGCTTTAAATTTGTGATGATACAGAAACC	5783
QY	6051	AGGCTTTGGGATACAAAGTTCTTTCTCTTCAATTTGATGCTGCGTGCACCTGTGTAAGACAGAT	6110
Db	5784	AGGCTTTGGGATACAAAGTTCTTTCTCTTCAATTTGATGCTGCGTGCACCTGTGTAAGACAGAT	5843
QY	6111	GTTTTTTGCGGAAATATAAAATTAATAGCTTGGAGCTTCGCGCAAAAAA	6163
Db	5844	GTTTTTTGCGGAAATATAAAATTAATAGCTTGGAGCTTCGCGCAAAAAA	5896

```

RESULT 3
AAH18243
ID AAH18243 standard; cDNA; 5573 BP.
XX
XX AAH18243;
AC
XX
XX 26-JUN-2001 (first entry)
DT
XX
XX Human cDNA sequence SEQ ID NO:18190.
DE
XX
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss
XX
XX Homo sapiens.
OS
XX
XX EP1074617-A2.
FN
XX
XX 07-FEB-2001.
PD
XX
XX 28-JUL-2000; 2000EP-0116126.
PF
XX
XX 29-JUL-1999; 99JP-0248036.
PR
XX 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
XX (HELI-) HELIX RES INST.
PA
XX
XX Ota T, Isogai T, Nishikawa T, Hayaishi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
XX WPI; 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602.
XX full-length cDNAs defined in the specification, and for the detection
XX and/or diagnosis of the abnormality of the proteins encoded by the
XX full-length cDNAs -
XX

```

XX Claim 8; SEQ ID 18190; 2537pp + CD ROM; English.

XX

CC The present invention describes primer sets for synthesising 5602

CC full-length cDNAs defined in the specification. Where a primer set

CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary

CC to the complementary strand of a polynucleotide which comprises one of

CC the 5602 nucleotide sequences defined in the specification, where the

CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination

CC of an oligonucleotide comprising a sequence complementary to the

CC complementary strand of a polynucleotide which comprises a 5'-end

CC sequence and an oligonucleotide comprising a sequence complementary to a

CC polynucleotide which comprises a 3'-end sequence, where the

CC oligonucleotide comprises at least 15 nucleotides and the combination of

CC the 5'-end sequence/3'-end sequence is selected from those defined in

CC the specification. The primer sets can be used in antisense therapy and

CC in gene therapy. The primers are useful for synthesising polynucleotides,

CC particularly full-length cDNAs. The primers are also useful for the

CC detection and/or diagnosis of the abnormality of the proteins encoded by

CC the full-length cDNAs. The primers allow obtaining of the full-length

CC cDNAs easily without any specialised methods. AAH03166 to AAH1628 and

CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to

CC AAB95893 represent human amino acid sequences; and AAH1629 to AAH16322

CC represent oligonucleotides, all of which are used in the exemplification

CC of the present invention.

XX

XX Sequence 5573 BP; 1373 A; 1524 C; 1290 G; 1386 T; 0 other;

XX

Query Match	77.7% ; Score 4786 ; DB 22 ; Length 5573 ;								
Best Local Similarity	99.8% ; Pred. No. 0 ;								
Matches 5566 ;	Conservative 0 ; Mismatches 6 ; Indels 4 ; Gaps 4 ;								
Qy	569 ATATGACTGTATGTCAGGCCCATGTCTCTTTGGGAGGTTCACTTCAGCCCTGCACAT	628							
Db	1 ATATTGACTGTATGTCAAGGCCCATGTCTCTTTGGGAGGTTCACTTCAGCCCTGCACAT	60							
Qy	629 CGTTGTTGCCACCCCTCAAGAACTTTCATCTCTGGGATGTCAAGCTCATTAAGGCATCG	688							
Db	61 CGTTGTTGCCACCCCTCAAGAACTTTCATCTCTGGGATGTCAAGGCATTAAGGCATCG	120							
Qy	689 GTTTAGAGCTGCAGATTTTTCATCCCTCGCCTGAGGAGATCGTCCGGGTGAGAGCTGCC	748							
Db	121 G-TTAGAGCTGCAGATTTTTCATCCCTCGCCTGAGGAGATCGTCCGGGTGAGAGCTGCC	179							
Qy	749 CAGACGGAGTCACTCATCTCCATCAGCGGCGGAATGCATGCCACCGTGTCAAGATCGGAA	808							
Db	180 CAGACGGAGTCACTCATCTCCATCAGCGGCGGAATGCATGCCACCGTGTCAAGATCGGAA	239							
Qy	809 CCTTCTGAGCAATGAGCACTGTGTCCGGATCAAGATGCAAGAGAGTAAAAATGAGCT	868							
Db	240 CCTTCTGAGCAATGAGCACTGTGTCCGGATCAAGATGCAAGAGAGTAAAAATGAGCT	299							
Qy	869 TACACCTTCCAGGTTTCAACCCAGAAATGTCTCCGGCTTCAGCATTTGCCAAACCGCTCAT	928							
Db	300 TACACCTTCCAGGTTTCAACCCAGAAATGTCTCCGGCTTCAGCATTTGCCAAACCGCTCAT	359							
Qy	929 CTATTAACAGCTGTGCATCATGAGTCTGTGTTGAGGGTGAAGGCTCAGCAACCCCTGA	988							
Db	360 CTATTAACAGCTGTGTGCATCATGAGTCTGTGTTGAGGGTGAAGGCTCAGCAACCCCTGA	419							
Qy	989 TGTCTGCCAATCAACCCAGAAAGCTTCCCTGAGATGAGCTCATGACGTGAGATTGTTCG	1048							
Db	420 TGTCTGCCAATCAACCCAGAAAGCTTCCCTGAGATGAGCTCATGACGTGAGATTGTTCG	479							
Qy	1049 TTCTCTGCACACTTCGCGGACAGAGCTTCTCTTCTTCAACTTCAACTCTTCCAATCTGAGA	1108							
Db	480 TTCTCTGCACACTTCGCGGACAGAGCTTCTCTTCTTCAACTTCAACTCTTCCAATCTGAGA	539							
Qy	1109 GGAAGAGAGAGCGGGTTGAATATCAATATCCCGGAGCTCCACACCAACCCCGAGGTGTCCA	1168							
Db	540 GGAAGAGAGAGCGGGTTGAATATCAATATCCCGGAGCTCCACACCAACCCCGAGGTGTCCA	599							
Qy	1169 AGCTGAGAGACAGACGCTGGGAACATGCGGGGAACCTTCAACCTCTCTTCGAAAGCT	1228							



Db 600 AGTGGAGGACGAGCCCTGGGAGCATGGCGGGAACTTCAACCTCTCTGCAAGGCT 659  
Qy 1229 GTGACCAAGATGCGCCAAAGTCCAGGGATCTCCGGCTGCAAGTTTCCAAAGTTTGGTCCAA 1288  
Db 660 GTGACCAAGATGCGCCAAAGTCCAGGGATCTCCGGCTGCAAGTTTCCAAAGTTTGGTCCAA 719  
Qy 1289 ATTCACAAAATGAAAGCAATAAATCTAGCTGGTTGACTTGAATGAAGCAAGCCATGT 1348  
Db 720 ATTCACAAAATGAAAGCAATAAATCTAGCTGGTTGACTTGAATGAAGCAAGCCATGT 779  
Qy 1349 CACTCAACATTCAGGCCAGCGGCCGTGTAACAGAGCCGCAAGTTTGTCCCTGGCTGTTTGG 1408  
Db 780 CACTCAACATTCAGGCCAGCGGCCGTGTAACAGAGCCGCAAGTTTGTCCCTGGCTGTTTGG 839  
Qy 1409 TGTGTATGAATCTCGGACCTGAGTAGCAAGCACTGACCTGACATCTGGCTCCAAACACA 1468  
Db 840 TGTGTATGAATCTCGGACCTGAGTAGCAAGCACTGACCTGACATCTGGCTCCAAACACA 899  
Qy 1469 AAATCTCTCTCTTGTGTATGATCTGACAGCTCTGTGATGAAATGTGAAAAAACATTA 1528  
Db 900 AAATCTCTCTCTTGTGTATGATCTGACAGCTCTGTGATGAAATGTGAAAAAACATTA 959  
Qy 1529 GCTGCACAGACCAACCGGTACTGCAAGAAATCTTACTGACTCCAGGTGCCAGTAGACA 1588  
Db 960 GCTGCACAGACCAACCGGTACTGCAAGAAATCTTACTGACTCCAGGTGCCAGTAGACA 1019  
Qy 1589 TCCTCCACCTGCTGAGAGCTGAGTAGATCTTCTCCGGAAGCTGCTGGTCCCAAGGACA 1648  
Db 1020 TCCTCCACCTGCTGAGAGCTGAGTAGATCTTCTCCGGAAGCTGCTGGTCCCAAGGACA 1079  
Qy 1649 GGCTCAGCCTGTGTGTGTGCTGCCAGCCAGAAAGCTGAGAGATACACAGAAAGCCCT 1708  
Db 1080 GGCTCAGCCTGTGTGTGTGCTGCCAGCCAGAAAGCTGAGAGATACACAGAAAGCCCT 1139  
Qy 1709 GCACACCAAGCTTCAAGTACTCTGTGGCCAGTGCATCCAGCCAGAGACTGTACTTGG 1768  
Db 1140 GCACACCAAGCTTCAAGTACTCTGTGGCCAGTGCATCCAGCCAGAGACTGTACTTGG 1199  
Qy 1769 GCTCCTCTGCGCCGGAGGCTGTATCAAGCAGATCCAGGTGAAGCAAGCAATCTCGGTGA 1828  
Db 1200 GCTCCTCTGCGCCGGAGGCTGTATCAAGCAGATCCAGGTGAAGCAAGCAATCTCGGTGA 1259  
Qy 1829 CCCTTTCGCACTTTGGCCCGCAGCTTCCAAAGAGGCTCCAGGCGAGGTCTGACGGTGT 1888  
Db 1260 CCCTTTCGCACTTTGGCCCGCAGCTTCCAAAGAGGCTCCAGGCGAGGTCTGACGGTGT 1319  
Qy 1889 CCTTTATACCTTATTTCAAGAGAAAGCGGTTTTCACGGTGAACCCCTGACACAAAAAGCA 1948  
Db 1320 CCTTTATACCTTATTTCAAGAGAAAGCGGTTTTCACGGTGAACCCCTGACACAAAAAGCA 1379  
Qy 1949 AGGTCTACCTGAGAGACCCCAACTGGGACCGGGGCTGCAATCCCTCACTGTGTCTT 2008  
Db 1380 AGGTCTACCTGAGAGACCCCAACTGGGACCGGGGCTGCAATCCCTCACTGTGTCTT 1439  
Qy 2009 GGAACATCAAGCTGCGCCAGAGACAGGTGAGCTGCTGACTTTCTTTAAGAGCGGAGCG 2068  
Db 1440 GGAACATCAAGCTGCGCCAGAGACAGGTGAGCTGCTGCTGACTTTCTTTAAGAGCGGAGCG 1499  
Qy 2069 GCGTGGTCTGCGAGACGGGCGCGCATTCATGATCATCCAGAGACAGCGGAGCG 2128  
Db 1500 GCGTGGTCTGCGAGACGGGCGCGCATTCATGATCATCCAGAGACAGCGGAGCG 1559  
Qy 2129 AGGAATCTTCAGCCCTGAGCGAGATGTGTCTCCCAAGCAAGTTTCCACATCACAGCT 2188  
Db 1550 AGGAATCTTCAGCCCTGAGCGAGATGTGTCTCCCAAGCAAGTTTCCACATCACAGCT 1619  
Qy 2189 TCTGGGTCAACATCTCTAAGTCCAGCCCAAGAGCGGCAAGGAGCTAGAGCTGTCTTCT 2248  
Db 1620 TCTGGGTCAACATCTCTAAGTCCAGCCCAAGAGCGGCAAGGAGCTAGAGCTGTCTTCT 1679  
Qy 2249 CGGTGACATTTACCCCAAGAGCTGTGACTTGAATCTGTCAATCCGACGCGGTGGAG 2308

Db 1680 CGGTGACATTTACCCCAAGAGCTGTGACTGTGACTGTCAATCTGTGACAGCGGTGGAG 1739  
Qy 2309 GTGAGTCTTACTGTCTGTCTGCTCGGCTCGGAGCTCATCATTTGTGTGAAAAAGAAATA 2368  
Db 1740 GTGAGTCTTACTGTCTGTCTGCTCGGCTCGGAGCTCATCATTTGTGTGAAAAAGAAATA 1799  
Qy 2369 AGAAGCAAAACAAAGGCCCCCGTGTGGTATCTAACATGGCAACATCAATCTGAGATGC 2428  
Db 1800 AGAAGCAAAACAAAGGCCCCCGTGTGGTATCTAACATGGCAACATCAATCTGAGATGC 1859  
Qy 2429 CGAGGACGCAAAAAAGTTTTCAGAAAGGCGCAAGCAATGATCTCCATGTGTATGCA 2488  
Db 1860 CGAGGACGCAAAAAAGTTTTCAGAAAGGCGCAAGCAATGATCTCCATGTGTATGCA 1919  
Qy 2489 TCATGAGGACACCATGTGATATGAGGCAATGTGCTACAGGATTCAGCGGCTCTTCCTGC 2548  
Db 1920 TCATGAGGACACCATGTGATATGAGGCAATGTGCTACAGGATTCAGCGGCTCTTCCTGC 1979  
Qy 2549 AGCCAGAGTGGACACCTACCGGCGGTTCCAGGACCATAGGGGCTGTCTCCCTCC 2608  
Db 1980 AGCCAGAGTGGACACCTACCGGCGGTTCCAGGACCATAGGGGCTGTCTCCCTCC 2039  
Qy 2609 CACCCACATATGTCTTCAGGGCCCCCACTGCAAAAGTTGGCCACTGAGAGCCACTCTC 2668  
Db 2040 CACCCACATATGTCTTCAGGGCCCCCACTGCAAAAGTTGGCCACTGAGAGCCACTCTC 2099  
Qy 2669 GCTCCCTCTCGAGTCTGAGTGAACGGTACACCTTCTCCATCCCAACATGGGGATG 2728  
Db 2100 GCTCCCTCTCGAGTCTGAGTGAACGGTACACCTTCTCCATCCCAACATGGGGATG 2159  
Qy 2729 TAAGCAGAAAGACACAGACATTCCTTACTGAACTCAGAGGCCATGAGCCAGAG 2788  
Db 2160 TAAGCAGAAAGACACAGACATTCCTTACTGAACTCAGAGGCCATGAGCCAGAG 2219  
Qy 2789 AATTAATTTGATTCATTCAGAGCTTTGTGATTTCAATTAAGCAGGCACTGAGACAC 2848  
Db 2220 AATTAATTTGATTCATTCAGAGCTTTGTGATTTCAATTAAGCAGGCACTGAGACAC 2279  
Qy 2849 CGTCCGTCTCTTACCAAGAAATCTTAAGAAAGAAATTAAGAAAGAAACACAGCA 2908  
Db 2280 CGTCCGTCTCTTACCAAGAAATCTTAAGAAAGAAATTAAGAAAGAAACACAGCA 2339  
Qy 2909 GGTTCCTCGACACCGCCCACTTCACTGAGTGTCTGAGTCACTTCAAGGCAAGACA 2968  
Db 2340 GGTTCCTCGACACCGCCCACTTCACTGAGTGTGTGATGATCTTCAAGGCAAGACA 2399  
Qy 2969 TTGAATAATGATGAATTCGAATCTGAGTACAGTCAATGATGATGATGATGATGATGAT 3028  
Db 2400 TTGAATAATGATGAATTCGAATCTGAGTACAGTCAATGATGATGATGATGATGATGAT 2459  
Qy 3029 GGTTCGTGGTTAGGCAAGCTGTATATGAGAGAGAGAGGCTGTGATCACTTACGCTAAGG 3088  
Db 2460 GGTTCGTGGTTAGGCAAGCTGTATATGAGAGAGAGAGGCTGTGATCACTTACGCTAAGG 2519  
Qy 3089 TTGACAGAAAGCTTCAAGAGTGTAAACAGAGGCTTCCCTCTTCAAGCAACAGT 3148  
Db 2520 TTGACAGAAAGCTTCAAGAGTGTAAACAGAGGCTTCCCTCTTCAAGCAACAGT 2579  
Qy 3149 TCCAAATTCGAAGAGCTTACCTGAGTCCCTCACTGAGGAGTCCCAAGATGAATAAC 3208  
Db 2580 TCCAAATTCGAAGAGCTTACCTGAGTCCCTCACTGAGGAGTCCCAAGATGAATAAC 2639  
Qy 3209 GACAAATGTGCTTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 3268  
Db 2640 GACAAATGTGCTTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 2699  
Qy 3269 GTATAACCACTAGCTCTTTTCACTGACTTATTAATTAATCAATCTAATCTAGTGTGAT 3328  
Db 2700 GTATAACCACTAGCTCTTTTCACTGACTTATTAATTAATCAATCTAATCTAGTGTGAT 2759  
Qy 3329 GCTTGGTGTGTGACTTCTACTGACCGCTGATTAACAGTGTGCTGTCCCAAGGTGTGG 3388  
Db 2760 GCTTGGTGTGTGACTTCTACTGACCGCTGATTAACAGTGTGCTGTCCCAAGGTGTGG 2819

QY	3389	GAATATATTACAACTGTGCAACCAAGAAATGTCGTGTTTGGACGATTTGACA	3448
Db	2820	GAATATATTACAACTGTGCAACCAAGAAATGTCGTGTTTGGACGATTTGACA	2879
QY	3449	TATCTGCTTTATATAGACATTCCTGATCTCTAGTGGTTCGTGTTATCCATTGTG	3508
Db	2880	TATCTGCTTTATATAGACATTCCTGATCTCTAGTGGTTCGTGTTATCCATTGTG	2939
QY	3509	GAATATCATCTTGAATCCCATTTGTCCTATAGCTTACAGAAATTTCTCAAGT	3568
Db	2940	GAATATCATCTTGAATCCCATTTGTCCTATAGCTTACAGAAATTTCTCAAGT	2999
QY	3569	TTCCATGTGCGGTTCTCCTAGCTGAGCAATCTTGACATTTAAAGAAATTTAGAG	3628
Db	3000	TTCCATGTGCGGTTCTCCTAGCTGAGCAATCTTGACATTTAAAGAAATTTAGAG	3059
QY	3629	ATATTTTCATCCTCTAAAAATGTTTAAATATATACCAACAGTGCCCTCTGATAGT	3688
Db	3060	ATATTTTCATCCTCTAAAAATGTTTAAATATATACCAACAGTGCCCTCTGATAGT	3119
QY	3689	TTCTGTGGCCACTGCAACCCATTCCTGTAGCTTTAAACAACAATTAACCTATAGTC	3748
Db	3120	TTCTGTGGCCACTGCAACCCATTCCTGTAGCTTTAAACAACAATTAACCTATAGTC	3179
QY	3749	CTGGGGATCAGAAATTCCAAAAATGATGTCCCTGATGAAAAATCAAGGTGACAGAGCT	3808
Db	3180	CTGGGGATCAGAAATTCCAAAAATGATGTCCCTGATGAAAAATCAAGGTGACAGAGCT	3239
QY	3809	GTGCTCCTCTTGAAGGCTCTAGGGAGAACCGGTTCTTGCCATTTCAAGCTTCTAGAG	3868
Db	3240	GTGCTCCTCTTGAAGGCTCTAGGGAGAACCGGTTCTTGCCATTTCAAGCTTCTAGAG	3299
QY	3869	CTGGCTGATTTCCCAAGGCTCCAGTGGCTGTCAAGCTTTCTCATGAGCATCACTGGA	3928
Db	3300	CTGGCTGATTTCCCAAGGCTCCAGTGGCTGTCAAGCTTTCTCATGAGCATCACTGGA	3359
QY	3929	CATGCGCCCTCCCACTTCCCTCTTTGACTTACAAAGCCACACAGAGATCCAGATAT	3988
Db	3360	CATGCGCCCTCCCACTTCCCTCTTTGACTTACAAAGCCACACAGAGATCCAGATAT	3419
QY	3989	CTCTCCATCTTAAAGTCTCTTATCATCTCTGGAAGGCTTTTGCATGCAAGCAACATA	4048
Db	3420	CTCTCCATCTTAAAGTCTCTTATCATCTCTGGAAGGCTTTTGCATGCAAGCAACATA	3479
QY	4049	GCCACAGGTGGGATTTAGAACACAGACATCTTTGGGTCGTATTCTGTGCTACACAC	4108
Db	3480	GCCACAGGTGGGATTTAGAACACAGACATCTTTGGGTCGTATTCTGTGCTACACAC	3539
QY	4109	CTTCTGTGCACTGACTCCCAACAGAGAGGCTTACAAATGATCTGGCGCACAGGATGTTT	4168
Db	3540	CTTCTGTGCACTGACTCCCAACAGAGAGGCTTACAAATGATCTGGCGCACAGGATGTTT	3599
QY	4169	TGTTTAGCTTCCGGACTTAACACTT-AAAAAAACCAGATTCAGAAATCTGGCGATGC	4227
Db	3600	TGTTTAGCTTCCGGACTTAACACTTAAAAAAAACCCAGATTCAGAAATCTGGCGATGC	3659
QY	4228	TGGGGCTCAATCTCACTAGCAACAATCGTGCTGAGCTGCGCACACAGCTCTCGCTTTA	4287
Db	3660	TGGGGCTCAATCTCACTAGCAACAATCGTGCTGAGCTGCGCACACAGCTCTCGCTTTA	3719
QY	4288	GAAAGGGGTGTCACCTTACACAGGTTCACACAGGCCACACTAGGCCCTATACCTTCCACA	4347
Db	3720	GAAAGGGGTGTCACCTTACACAGGTTCACACAGGCCACACTAGGCCCTATACCTTCCACA	3779
QY	4348	ATGAGGCTAAGTGTGTTGTTTCTACTGATCAATGCCCCCTGACAGGTTGCAATTATGTAATG	4407
Db	3780	ATGAGGCTAAGTGTGTTGTTTCTACTGATCAATGCCCCCTGACAGGTTGCAATTATGTAATG	3839
QY	4408	AAAAAGAAAGACTGGGATTAATCTTAAATCAGGTAGTAAACATAGAACAAATGTGTGC	4467
Db	3840	AAAAAGAAAGACTGGGATTAATCTTAAATCAGGTAGTAAACATAGAACAAATGTGTGC	3899

[illegible]



```

Db      4978 TTTAAAGCCAGGCCCCCACTTACATTAACATTTACAGGCTGACCTGAAACAGTGAAGTGC
Oy      5608 CATTGTTGAGGCTACTGATCCAGCCAGCTGCTCATCCAGTGTGCGCAGCCCTA
Db      5038 CATTGTTGAGGCTACTGATCCAGCCAGCTGCTCATCCAGTGTGCGCAGCCCTA
Oy      5668 CGAGGAAGCCAGCCAGTCCAGGAGCTGTCTCTAATGCTGTGCTATTGACAGAGAGGA
Db      5098 CGAGGAAGCCAGCCAGTCCAGGAGCTGTCTCTAATGCTGTGCTATTGACAGAGAGGA
Oy      5728 AAGGTTCTAAGAGAGAGTCACTGGGACAAAGCCACCGGACATGCGCTTGGTAA
Db      5158 AAGGTTCTAAGAGAGAGTCACTGGGACAAAGCCACCGGACATGCGCTTGGTAA
Oy      5788 AAGTTAGAGAGCTGCTGTGTGGATCTGACAGTCTCACTGGAATTAATTTATTCATTG
Db      5218 AAGTTAGAGAGCTGCTGTGTGGATCTGACAGTCTCACTGGAATTAATTTATTCATTG
Oy      5848 CAGATACCTTTTAAAGTGGCATTTTATTCATTTCTGTGCTTTAAATTAACAATGACCA
Db      5278 CAGATACCTTTTAAAGTGGCATTTTATTCATTTCTGTGCTTTAAATTAACAATGACCA
Oy      5908 AAAAAACAGTATCAGCTGTTTAAAGTCTTCCGCTACTTGTCCCTGTTCAAGTAGAGC
Db      5338 AAAAAACAGTATCAGCTGTTTAAAGTCTTCCGCTACTTGTCCCTGTTCAAGTAGAGC
Oy      5968 CCGGTTTCCAGTGTGTTGATCTGTCAGGCTCAGGATGGGCTCAGCAATCTGCTTA
Db      5398 CCGGTTTCCAGTGTGTTGATCTGTCAGGCTCAGGATGGGCTCAGCAATCTGCTTA
Oy      6028 ATTGTGATGATATACAGAAAGCCAGCTTGGGATACAGTTCTTCTCTTCATTGAT
Db      5458 ATTGTGATGATATACAGAAAGCCAGCTTGGGATACAGTTCTTCTCTTCATTGAT
Oy      6088 GCCGTGCTGCTGTGTGAAGCAGATGTTTTTTCGGAATTAATAATATAGCTTGG
Db      5518 GCCGTGCTGCTGTGTGAAGCAGATGTTTTTTCGGAATTAATAATATAGCTTGG

```

RESULT 4  
AB211770  
ID AB211770 standard; cDNA; 2669 BP.

AC AB211770;  
XX  
XX  
DT 20-JAN-2003 (first entry)  
XX  
XX  
DE Human polynucleotide SEQ ID NO 652.  
XX  
XX  
KM Human; genome mapping; gene therapy; food supplement; virus; fungus;  
KM cell-proliferative disorder; neurodegenerative disease; bacterial;  
KM Parkinson's disease; Alzheimer's disease; autoimmune disease;  
KM multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;  
KM arthritis; cystic fibrosis; immunomodulator; neoplastic; dermatological;  
KM antiparkinsonian; antidiabetic; immunosuppressive; neuroprotection;  
KM haemostatic; vulnary; fungicide; antibacterial; virucide; protozoacide;  
KM antiaerobic; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200270539-A2.  
XX  
PD 12-SEP-2002.  
XX  
XX  
PF 05-MAR-2002; 2002WO-US05095.  
XX  
PR 05-MAR-2001; 2001US-0799451.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;  
Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;

```

PI Wehman T, Wang J, Wang D, Drmanac RT;
XX WPI; 2002-759812/82.
DR P-PSDB; ABP69553.
XX
PT New polynucleotides comprising sequences assembled from expressed
PT sequence tags (ESTs), useful for treating cell-proliferative,
PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or
PT platelet or coagulation disorders
XX
PS Claim 1; SEQ ID NO 652; 1012pp + Sequence Listing; English.
XX
XX The invention relates to an isolated polynucleotide (1) comprising a
CC nucleotide sequence selected from any of 948 sequences
CC (AB21119-AB212066) or their mature protein coding portion, active domain
CC coding protein or complementary sequences. The polynucleotides are useful
CC for identifying expressed genes or for physical mapping of human genome.
CC The encoded polypeptides (ABP6902-ABP6949) are useful as molecular
CC weight markers, as a food supplement, for generating antibodies, in
CC medical imaging, screening and diagnostic assays and for treating
CC cell-proliferative disorders (cancer), neurodegenerative diseases
CC (Parkinson's or Alzheimer's disease), autoimmune diseases (multiple
CC sclerosis, diabetes, lupus) genetic disorders, myeloid or lymphoid
CC disorders, platelet or coagulation disorders, wound, burns, incision,
CC ulcers, liver or lung fibrosis, infections (bacterial, viral, fungal,
CC parasitic), arthritis, etc.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2669 BP; 634 A; 790 C; 675 G; 570 T; 0 other;

```

Query Match 41.0%; Score 2524; DB 24; Length 2669;  
Best Local Similarity 99.9%; Pred. No. 0; Mismatches 2; Indels 0; Gaps 0;  
Matches 2624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

Oy      219 GGGCGCTGAGGCGCGCGCTCCCAACGTCGTTTCCCAACGAGCGGAGCGTCCCG
Db      44 GGGCGCTGAGGCGCGCGCTCCCAACGTCGTTTCCCAACGAGCGGAGCGTCCCG
Oy      279 AGTCATGCGCGGCTGAACTGCGGCGCTCTATGCGACTGCTAGGGCTTGTGCTGCG
Db      104 AGTCATGCGCGGCTGAACTGCGGCGCTCTATGCGACTGCTAGGGCTTGTGCTGCG
Oy      339 TGGCGGCGGCGCGCGCGGCGGCGAGAGCTTTGAGATGCTGCGCAGCGAAGCA
Db      164 TGGCGGCGGCGCGCGCGGCGGCGAGAGCTTTGAGATGCTGCGCAGCGAAGCA
Oy      399 CATTACAGTCTCATTAAGCTGGGAGCCCGACTGCTGCGCAAAACCTGTTACATGCT
Db      224 CATTACAGTCTCATTAAGCTGGGAGCCCGACTGCTGCGCAAAACCTGTTACATGCT
Oy      459 CATTCTTAAGAAGCATATACCATGTTGTCATCAAGTCTGGAGAAAGATAGCTTTAC
Db      284 CATTCTTAAGAAGCATATACCATGTTGTCATCAAGTCTGGAGAAAGATAGCTTTAC
Oy      519 CTTTACGTCGAGAGTCTGTGAATCACTTTGTCTATAGATCCGAAATATATGACTG
Db      344 CTTTACGTCGAGAGTCTGTGAATCACTTTGTCTATAGATCCGAAATATATGACTG
Oy      579 TATGTCAGGCGCATGCTTTTGGGAGGTTGAGCTTACAGCTTACAGCTTGTGCTCC
Db      404 TATGTCAGGCGCATGCTTTTGGGAGGTTGAGCTTACAGCTTGTGCTTGTGCTCC
Oy      639 TACCTCAACAGAACTTTTCACTGAGATGTCAAGCTCATTAAGAGCATGCTTTAGACT
Db      464 TACCTCAACAGAACTTTTCACTGAGATGTCAAGCTCATTAAGAGCATGCTTTAGACT
Oy      699 GCAGTTTTCATCCCTGCGCTGAGGAGATGCTGCTGCGGCTGAGAGCTCCAGAGAGGT
Db      524 GCAGTTTTCATCCCTGCGCTGAGGAGATGCTGCTGCGGCTGAGAGCTCCAGAGAGGT
Oy      759 CACTCATTCCATCAGCGGCGCAATGATGACACCGTGTGAGATCGGAACCTTCTGAG

```

Db 584 CACTCCTCATAGAGGCGGAATGATGACCGGTGTCAGGATGGAACCTTTCGACG 643  
Qy 819 CAATGCACTGTGTCCCGATCAAGATGCAAGAAGATGAAATAGCTTACCTCC 878  
Db 644 CAATGCACTGTGTCCCGATCAAGATGCAAGAAGATGAAATAGCTTACCTCC 703  
Qy 879 ATGGTTCCACCCAGAAATGTCTCCGCTTACGATTCGAAACCGCTCATCTATAAAG 938  
Db 704 ATGGTTCCACCCAGAAATGTCTCCGCTTACGATTCGAAACCGCTCATCTATAAAG 763  
Qy 939 TCTGTGATCATGAGCTGTGTGTTGAGGGTGAAGGTGAGCAACCTGATGTCTGCAA 998  
Db 764 TCTGTGATCATGAGCTGTGTGTTGAGGGTGAAGGTGAGCAACCTGATGTCTGCAA 823  
Qy 999 CTACCCAGAAAGCTTCCCTGAGATGAGCTCATGAGTGCAGTTTGTCTGTCACCA 1058  
Db 824 CTACCCAGAAAGCTTCCCTGAGATGAGCTCATGAGTGCAGTTTGTCTGTCACCA 883  
Qy 1059 CTTGCGGGCAGCGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1118  
Db 884 CTTGCGGGCAGCGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 943  
Qy 1119 GCGGGTTGAATATACATCCCGGGCTCCACCAACCCGAGGTGTTCAAGCTGAGAGA 1178  
Db 944 GCGGGTTGAATATACATCCCGGGCTCCACCAACCCGAGGTGTTCAAGCTGAGAGA 1003  
Qy 1179 CAAGCAAGCTGGGAACATGCGGGGAACTTCAACTCTCTCTGCAAGGCTGTGACCAAG 1238  
Db 1004 CAAGCAAGCTGGGAACATGCGGGGAACTTCAACTCTCTCTGCAAGGCTGTGACCAAG 1063  
Qy 1239 TGCCCAAGTCCAGGATCTCTCGGCTGCAAGTTCCAAAGTTTGTGTCACATCCAAAA 1298  
Db 1064 TGCCCAAGTCCAGGATCTCTCGGCTGCAAGTTTGTGTCACATCCAAAA 1123  
Qy 1299 TGAAGAAATAAATCTACGTGTGTTGACTGAGTAAGACGAGCCATGCACTCAACAT 1358  
Db 1124 TGAAGAAATAAATCTACGTGTGTTGACTGAGTAAGACGAGCCATGCACTCAACAT 1183  
Qy 1359 CGAGCCACGCGCCGCTCAACAGAGCGCAAGTTGTCTGAGTGTCTGATGA 1418  
Db 1184 CGAGCCACGCGCCGCTCAACAGAGCGCAAGTTGTCTGAGTGTCTGATGA 1243  
Qy 1419 ATCTCGGACTGCAATGCAACCTCACTGCAATCTGCTCCAAACAAAAATCTCTT 1478  
Db 1244 ATCTCGGACTGCAATGCAACCTCACTGCAATCTGCTCCAAACAAAAATCTCTT 1303  
Qy 1479 CCTTTGATGATCTGACAGCTCTGTGAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 1538  
Db 1304 CCTTTGATGATCTGACAGCTCTGTGAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 1363  
Qy 1539 CCAACGGTATGCAAGAAATCTTACTCACTCCAGGTGCCAGTGAACATCTCCACCT 1598  
Db 1364 CCAACGGTATGCAAGAAATCTTACTCACTCCAGGTGCCAGTGAACATCTCCACCT 1423  
Qy 1599 GCGTGTGAGCTGATGATCTTCTTGTGAAGCTGTGTGTCCTCAAGACAAGCTCACT 1658  
Db 1424 GCGTGTGAGCTGATGATCTTCTTGTGAAGCTGTGTGTCCTCAAGACAAGCTCACT 1483  
Qy 1659 GGTGTGTGTGTCAGCCCAAGAAAGTGTGAGAGCATACAGAAAGAAAGCTGTGACAAG 1718  
Db 1484 GGTGTGTGTGTCAGCCCAAGAAAGTGTGAGAGCATACAGAAAGAAAGCTGTGACAAG 1543  
Qy 1719 CTTCAAGTACTGCTGTGAGCATATCCAGCAGAGACCTGTACTTGGGCTCTTCTG 1778  
Db 1544 CTTCAAGTACTGCTGTGAGCATATCCAGCAGAGACCTGTACTTGGGCTCTTCTG 1603  
Qy 1779 CCGGGAAGGCTCATCAAGCATCCAGTGAAGCAAAATCTCGGTGACCTTTCGAC 1838  
Db 1604 CCGGGAAGGCTCATCAAGCATCCAGTGAAGCAAAATCTCGGTGACCTTTCGAC 1663  
Qy 1839 CTTTGGCCCAAGCTTCCAAGAGGCTTCAAGCAGGCTGTGACGCTCTTTATACC 1898

Db 1664 CTTTGGCCCAAGCTTCCAAGAGGCTTCAAGCAGGCTGTGACGCTGTCTTTATACC 1723  
Qy 1899 TTAATTTCAAGAGGAAGGCGTTTTCACGATGACCCCTGACACAAAGAGCATACCT 1958  
Db 1724 TTAATTTCAAGAGGAAGGCGTTTTCACGATGACCCCTGACACAAAGAGCATACCT 1783  
Qy 1959 GAGAACCCCAACTGGGACCCGGGCTGTGATCTTCTGATCTTGTCTGTGAATATCG 2018  
Db 1784 GAGAACCCCAACTGGGACCCGGGCTGTGATCTTCTGATCTTGTCTGTGAATATCG 1843  
Qy 2019 CGTGCCCAAGAGCAAGTGTGCTGCTGCTGCTTCTTTAAGAGGCGAGGCGGTGTG 2078  
Db 1844 CGTGCCCAAGAGCAAGTGTGCTGCTGCTGCTTCTTTAAGAGGCGAGGCGGTGTG 1903  
Qy 2079 CCAGAAGGCGCGATTCATGATCATCAGAGAGCGGACCCGGCTGAGAGATCTT 2138  
Db 1904 CCAGAAGGCGCGATTCATGATCATCAGAGAGCGGACCCGGCTGAGAGATCTT 1963  
Qy 2139 CAGCTGAGCAGAGATGTGCTCCCAAGCAAGCTTCCACATCAAGCTTCTGGGTCAA 2198  
Db 1964 CAGCTGAGCAGAGATGTGCTCCCAAGCAAGCTTCCACATCAAGCTTCTGGGTCAA 2023  
Qy 2199 CATCTTAACTGACAGCCCAAGAGGCGCAAGAGCTAGACCTGCTTCTCGGTGACT 2258  
Db 2024 CATCTTAACTGACAGCCCAAGAGGCGCAAGAGCTAGACCTGCTTCTCGGTGACT 2083  
Qy 2259 TACCCCAAGAGCTGTGACTGTGATCTCTCATTCGACGCGTGTGAGTGTCTT 2318  
Db 2084 TACCCCAAGAGCTGTGACTGTGATCTCTCATTCGACGCGTGTGAGTGTCTT 2143  
Qy 2319 ACTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2378  
Db 2144 ACTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2203  
Qy 2379 CAAGGGCCCGCTGTGGGATCTAATAGGCAATCAATAGTGAATGAGTGCAGAGGCG 2438  
Db 2204 CAAGGGCCCGCTGTGGGATCTAATAGGCAATCAATAGTGAATGAGTGCAGAGGCG 2263  
Qy 2439 AAAAAAGTTTCAAGAAAGGCGAAGAGCAATGACTCCATGTGTATGCAATGAGGA 2498  
Db 2264 AAAAAAGTTTCAAGAAAGGCGAAGAGCAATGACTCCATGTGTATGCAATGAGGA 2323  
Qy 2499 CACCATGTATATGGGATCTGTCTACAGATTCAGAGGCTCTTCTGACCCAGAGGT 2558  
Db 2324 CACCATGTATATGGGATCTGTCTACAGATTCAGAGGCTCTTCTGACCCAGAGGT 2383  
Qy 2559 GGAACCTTACCGGCGCTTCAAGGCAACCATGGGGGTGTGTCTCTCCCTCCCAACCAT 2618  
Db 2384 GGAACCTTACCGGCGCTTCAAGGCAACCATGGGGGTGTGTCTCTCCCTCCCAACCAT 2443  
Qy 2619 ATGCTTCAAGGCGCCCAACTGCAAAAGTTGGCCACTGAGAGGACCACTCTGCTCCCTCC 2678  
Db 2444 ATGCTTCAAGGCGCCCAACTGCAAAAGTTGGCCACTGAGAGGACCACTCTGCTCCCTCC 2503  
Qy 2679 TGAATGTGAGTGAACCGTACACTTCTCCATTCCAAATGGGATGTAAAGCAAG 2738  
Db 2504 TGAATGTGAGTGAACCGTACACTTCTCCATTCCAAATGGGATGTAAAGCAAG 2563  
Qy 2739 GGAACAAGCATTTCCCTTACTGAACACTCAGGAGCCATGAGGCAAGATAACTTGA 2798  
Db 2564 GGAACAAGCATTTCCCTTACTGAACACTCAGGAGCCATGAGGCAAGATAACTTGA 2623  
Qy 2799 TCCATTTCAAGAGCTTGTGCTGAGTTTCAATAAGCAGGCACTGAGA 2844  
Db 2624 TCCATTTCAAGAGCTTGTGCTGAGTTTCAATAAGCAGGCACTGAGA 2669

RESULT 5  
AAA26351  
ID AAA26351 standard; cDNA; 2209 BP.  
XX  
AC  
XX AAA26351;

29-JUN-2000 (first entry)  
Human secreted protein gene 6 SEQ ID NO:16.  
Human; secreted protein; diagnosis; cytostatic; immunosuppressive;  
anti-HIV; antiinflammatory; neurotropic; neuroprotective; antiallergic;  
osteoporotic; antidiabetic; antibacterial; antidiabetic; antiashma;  
antiproliferative; cardiac; gene therapy; cancer; neurological disorder;  
immune disease; inflammation; blood disorder; tumour; ss.  
Homo sapiens.  
WO200006698-A1.  
10-FEB-2000.  
29-JUL-1999; 99WO-US17130.  
30-JUL-1998; 98US-0094657.  
05-AUG-1998; 98US-0095486.  
06-AUG-1998; 98US-0095454.  
06-AUG-1998; 98US-0095455.  
12-AUG-1998; 98US-0096319.  
(HUMA-) HUMAN GENOME SCI INC.  
Komatsu GA, Rosen CA, Ruben SM, Duan R, Moore PA, Shi Y,  
Lafleur D, Wei Y, Ni J, Florence KA, Young PE, Brewer LA,  
Soppet DR, Endres GA, Ebner R, Olsen HS, Mucenak M,  
WPI; 2000-195282/17.  
P-P8DB; AA91456.  
New isolated human genes and the secreted polypeptides they encode,  
useful for diagnosis and treatment of e.g. cancer, neurological  
disorders, immune diseases, inflammation or blood disorders -  
Claim 1; Page 375; 634pb; English.  
The polynucleotide sequences given in AA26346 to AA26458 encode the  
human secreted proteins given in AA91451 to AA91491. The human secreted  
proteins can have activities based on the tissues and cells they are  
expressed in. Examples of the activities are: cytostatic;  
immunosuppressive; anti-HIV; antiinflammatory; neurotropic; neuroprotective;  
antiallergic; osteoporotic; antidiabetic; antibacterial; antidiabetic;  
antiashma; antiproliferative; and cardiac. The polynucleotides and their  
corresponding secreted proteins are useful for preventing, treating or  
ameliorating medical conditions, e.g. by protein or gene therapy. Also  
pathological conditions can be diagnosed by determining the amount of the  
proteins in a sample or by determining the presence of mutations in the  
polynucleotides. Specific uses are described for each of the  
polynucleotides, based on which tissues they are most highly expressed  
in, and include developing products for the diagnosis or treatment of  
cancer, tumours, neurodegenerative disorders, developmental abnormalities  
and foetal deficiencies, blood disorders, diseases of the immune system,  
autoimmune diseases, hepatic and renal disorders, inflammation,  
allergies, Alzheimer's and behavioural disorders, schizophrenia,  
osteoporosis, arthritis, infections, AIDS, spinal cord injuries,  
transplant rejection, diabetes, asthma, sepsis, acne, psoriasis,  
cardiovascular disorders, reproductive disorders, gastrointestinal  
disorders, respiratory disorders and metabolic disorders. The proteins  
or polynucleotides can also be used as food additives or preservatives.  
The proteins are also useful for identifying their binding partners.  
AA26337 to AA26345 and AA91450 are sequences used in the  
exemplification of the present invention.  
Sequence 2209 BP; 508 A; 646 C; 566 G; 477 T; 12 other;  
Query Match 25.1%; Score 1545; DB 21; Length 2209;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 2145; Conservative 0; Mismatches 12; Indels 0; Gaps 0;  
204 AGCTGCGCGGCTTGGGCGCTGGGCGCGCTCCACCGTCTTTCCACCGAG 263

Db 2 AGCTGCGCGGCTTGGGCGCTGGGCGCGCTCCACCGTCTTTCCACCGAG 61  
264 GCGGAGGCGTCCCGAGCATGCGCGGCTGAATCGCGGGCTCTATGCACTGCTAG 323  
62 GCGGAGGCGTCCCGAGCATGCGCGGCTGAATCGCGGGCTCTATGCACTGCTAG 121  
324 GATTCTGCTGCTGGGCGCGGCTCCCGCGCGGCGGAGCAAGCTTTGAGATTGCTCT 383  
122 GATTCTGCTGCTGGGCGCGGCTCCCGCGCGGCGGAGCAAGCTTTGAGATTGCTCT 181  
384 GCGAGGAGGAGCAATTAAGTTCTCATTAAGCTGGGAGACCCGACTTGTGCGAAA 443  
182 GCGAGGAGGAGCAATTAAGTTCTCATTAAGCTGGGAGACCCGACTTGTGCGAAA 241  
444 ACCGCTTACATCGCATTTCTTAAAGCATATTAACATGTTGTCATCAAGTCTGAGA 503  
242 ACCGCTTACATCGCATTTCTTAAAGCATATTAACATGTTGTCATCAAGTCTGAGA 301  
504 AAGATAGCTCTTACCTTTAGCTGCGAGATCTGAGATCACTTGTATAGATGCA 563  
302 AAGATAGCTCTTACCTTTAGCTGCGAGATCTGAGATCACTTGTATAGATGCA 361  
564 GAAATATATGAATGATGATGAGGCGCATGCTCTTTGGGAGGTTCAAGCTTCAAGCTTC 623  
362 GAAATATATGAATGATGATGAGGCGCATGCTCTTTGGGAGGTTCAAGCTTCAAGCTTC 421  
624 GACATCGTTGTTGCTTACCTTACCTTCAAGAACTTTATCTGAGATGCAAGCTTAAAG 683  
422 GACATCGTTGTTGCTTACCTTCAAGAACTTTATCTGAGATGCAAGCTTAAAG 481  
684 CATCGTTTAAAGCTGAGTCTTTCATCCCTCGCTGAGGAGATCGGTCCGGGAGAG 743  
482 CATCGTTTAAAGCTGAGTCTTTCATCCCTCGCTGAGGAGATCGGTCCGGGAGAG 541  
744 CTGCGGAGGAGGAGTCACTCACTTCATGAGGCGGAGATGATGATGATGATGAT 803  
542 CTGCGGAGGAGGAGTCACTCACTTCATGAGGCGGAGATGATGATGATGATGAT 601  
804 CGGAACTTTTGCAGCATGAGTCTGCTCCCGATCAAGATGCAAGAGGATGAAT 863  
602 TGGAACTTTTGCAGCATGAGTCTGCTCCCGATCAAGATGCAAGAGGATGAAT 661  
864 GGCCTTACCTCCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 923  
662 GGCCTTACCTCCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 721  
924 CTGATCTTAAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 983  
722 CTGATCTTAAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 781  
984 CTTGATGCTGCTCAATCAAGGCTTCCCTGAGATGATGATGATGATGATGATGAT 1043  
782 CTTGATGCTGCTCAATCAAGGCTTCCCTGAGATGATGATGATGATGATGATGAT 841  
1044 TGTGCTTCTGCAACCTGCGGCGGCGGCTCTCTCTCAACTTCAACTTCAACTG 1103  
842 TGTGCTTCTGCAACCTGCGGCGGCGGCTCTCTCTCAACTTCAACTTCAACTG 901  
1104 TGAGAGGAGGAGGAGGCTTGAATATCAATCCCGGCTCAACCAACCCGAGGT 1163  
902 TRAGAGGAGGAGGAGGCTTGAATATCAATCCCGGCTCAACCAACCCGAGGT 961  
1164 GTTCAAGCTGAGGAGGAGGCTTGAATATCAATCCCGGCTCAACCAACCCGAGGT 1223  
962 GTTCAAGCTGAGGAGGAGGCTTGAATATCAATCCCGGCTCAACCAACCCGAGGT 1021  
1224 AGGCTGATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1283  
1022 AGGCTGATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1081  
1284 CCAATCATCAAAATGAAGCAATTAATCTAGTGTGATGATGATGATGATGATGAT 1343

[illegible]

CC tumour necrosis factor-alpha (TNF-alpha) from human blood, the  
CC proliferation or differentiation of chondrocytes, the proliferation or  
CC gene expression in pericyte cells, the release of proteoglycans from  
CC cartilage, the proliferation of inner ear utricular supporting cells or  
CC of T-lymphocytes, the release of a cytokine from peripheral blood  
CC monocytes (PBMCs), or the proliferation of endothelial cells. Some of  
CC the PRO polypeptides may modulate glucose or free fatty acid uptake by  
CC skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide  
CC to factor VIIa. The PRO polypeptides can be used in assays to identify  
CC molecules involved in binding interactions. The polynucleotides encoding  
CC PRO polypeptides can be used to generate probes, antisense RNA/DNA,  
CC transgenic or knock out animals and can be used in gene therapy.  
XX  
XX  
SQ Sequence 1376 BP; 311 A; 384 C; 352 G; 329 T; 0 other;

Query Match 18.1%; Score 1117; DB 22; Length 1376;  
Beet Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;  
Matches 1117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

190 CCAGGCGGAGCGCAGCTGCGCGGCTTGGGCGCTGGGGCGCGCTCCACCGCTG 249  
1 CCAGGCGGAGCGCAGCTGCGCGGCTTGGGCGCTGGGGCGCGCTCCACCGCTG 60  
250 TTTTCCCAACCGAGCGCGGCTCCGAGTCATGCGCGGCTGAACTGCGGGCTCT 309  
61 TTTTCCCAACCGAGCGCGGCTCCGAGTCATGCGCGGCTGAACTGCGGGCTCT 120  
310 ATCGCACTGCTAGGGGTTTCTGCTGCTGGGCGCGGCTCCGCGCGGCGGAGAGCT 369  
121 ATCGCACTGCTAGGGGTTTCTGCTGCTGGGCGCGGCTCCGCGCGGCGGAGAGCT 180  
370 TTGAGATTGCTCTGCGCAGAGAAAGCAATTACAGTTCTCATAAAGCTGGGGAACCCG 429  
181 TTGAGATTGCTCTGCGCAGAGAAAGCAATTACAGTTCTCATAAAGCTGGGGAACCCG 240  
430 ACTTGTGCGCAAAACCTGTTATCATGTCATTTCTAAAGACATTAACATGTTGTC 489  
241 ACTTGTGCGCAAAACCTGTTATCATGTCATTTCTAAAGACATTAACATGTTGTC 300  
490 ATCAAGTCTGGAAGAAAGATAGCTTTACTTTAGTGCAGAGTCTAGAAATCATCTT 549  
301 ATCAAGTCTGGAAGAAAGATAGCTTTACTTTAGTGCAGAGTCTAGAAATCATCTT 360  
550 GTCATAGAGATCCAGAAATAATTGACTGATGTCAGGCCCATGCTTTTGGGAGGTT 609  
361 GTCATAGAGATCCAGAAATAATTGACTGATGTCAGGCCCATGCTTTTGGGAGGTT 420  
610 CAGCTTACGCTCGACATGTTGTTGCTTACCTTCAACAGAACTTTCACTGGAGTTC 669  
421 CAGCTTACGCTCGACATGTTGTTGCTTACCTTCAACAGAACTTTCACTGGAGTTC 480  
670 AAAGCTCATTAAGAGATGCGTTTAAAGCTGAGCTTTTCACTCCCTGCTGAGGAGATC 729  
481 AAAGCTCATTAAGAGATGCGTTTAAAGCTGAGCTTTTCACTCCCTGCTGAGGAGATC 540  
730 GGTCCGGGAGAGAGCTGCGCAGAGGATCATCTCATCAGAGCGCGGAATTCATGTC 789  
541 GGTCCGGGAGAGAGCTGCGCAGAGGATCATCTCATCAGAGCGCGGAATTCATGTC 600  
790 ACCGTGTCAGAGATCGAACTTTCTGACAGCAATGGCACTGTGTCCCGATCAAGATGCAA 849  
601 ACCGTGTCAGAGATCGAACTTTCTGACAGCAATGGCACTGTGTCCCGATCAAGATGCAA 660  
850 GAAGAGTGAATAATGGCTTACCTCCATCCATGTTCCACCCAGAAATGTCCTCCGCTTC 909  
661 GAAGAGTGAATAATGGCTTACCTCCATCCATGTTCCACCCAGAAATGTCCTCCGCTTC 720  
910 ASCATTGCAAAACCGCTCATCTATAAAGCTGTGTCATCATGAGATCTGTTTGAAGGT 969  
721 AGCATTTGCAAAACCGCTCATCTATAAAGCTGTGTCATCATGAGATCTGTTTGAAGGT 780  
970 GAAGGCTCAGCAACCTGATGTCGCAACTAACAGAGGCTTCCCTGAGATGAGCTC 1029

DB 781 GAAGGCTCAGCAACCTGATGTCGCAACTAACAGAGGCTTCCCTGAGATGAGCTC 840  
QY 1030 ATGAGTGGCAGGTTTGTGCTTCTGACACCTGCGGGCGAGGCTCCTCTCAACTTC 1089  
DB 841 ATGAGTGGCAGGTTTGTGCTTCTGACACCTGCGGGCGAGGCTCCTCTCAACTTC 900  
QY 1090 AACCTTCCAACTGTAGAGAGAGAGAGAGCGGGTTGAATACATATCCCGGCTTCAAC 1149  
DB 901 AACCTTCCAACTGTAGAGAGAGAGAGAGCGGGTTGAATACATATCCCGGCTTCAAC 960  
QY 1150 ACCAACCCCGAGGCTTCAAGCTGAGAGCAACGAGCTGGGAAACATGCGGGAACTTC 1209  
DB 961 ACCAACCCCGAGGCTTCAAGCTGAGAGCAACGAGCTGGGAAACATGCGGGAACTTC 1020  
QY 1210 AACCTTCTCTGCAAGGCTGTGACCAAGTCCCAAGTCCGAGATCCTCGGCTGCAG 1269  
DB 1021 AACCTTCTCTGCAAGGCTGTGACCAAGTCCCAAGTCCGAGATCCTCGGCTGCAG 1080  
QY 1270 TTCCAAGTTTGTGTCACATCCACAAATGAAGCA 1306  
DB 1081 TTCCAAGTTTGTGTCACATCCACAAATGAAGCA 1117  
RESULT 7  
ID ACA03683  
XX ACA03683 strand; cDNA; 1376 BP.  
AC ACA03683;  
XX  
XX 23-MAY-2003 (first entry)  
DE  
XX cDNA encoding human PRO polypeptide #81.  
XX  
XX Human; PRO polypeptide; secreted and transmembrane protein;  
KW Tumour necrosis factor-alpha; TNF-alpha; blood; proliferation;  
KW cytosolic; gene; ss.  
XX  
OS Homo sapiens.  
XX  
XX US2003036180-A1.  
PN  
XX  
PD 20-FEB-2003.  
PF  
XX  
PF 09-MAY-2002; 2002US-0143114.  
XX  
XX 31-MAR-1997; 97WO-US05230.  
PR 12-JUN-1998; 98WO-US12456.  
PR 14-JUN-1998; 98WO-US14552.  
PR 28-AUG-1998; 98WO-US17888.  
PR 10-SEP-1998; 98WO-US18824.  
PR 14-SEP-1998; 98WO-US19093.  
PR 14-SEP-1998; 98WO-US19094.  
PR 14-SEP-1998; 98WO-US19177.  
PR 16-SEP-1998; 98WO-US19330.  
PR 17-SEP-1998; 98WO-US19437.  
PR 07-OCT-1998; 98WO-US21141.  
PR 29-OCT-1998; 98WO-US22991.  
PR 29-OCT-1998; 98WO-US22992.  
PR 20-NOV-1998; 98WO-US24855.  
PR 01-DEC-1998; 98WO-US25108.  
PR 05-JAN-1999; 99WO-US00106.  
PR 08-MAR-1999; 99WO-US05028.  
PR 10-MAR-1999; 99WO-US05190.  
PR 20-APR-1999; 99WO-US08615.  
PR 14-MAY-1999; 99WO-US10733.  
PR 02-JUN-1999; 99WO-US12252.  
PR 01-SEP-1999; 99WO-US20111.  
PR 08-SEP-1999; 99WO-US20594.  
PR 13-SEP-1999; 99WO-US20944.  
PR 15-SEP-1999; 99WO-US21090.  
PR 15-SEP-1999; 99WO-US21547.  
PR 05-OCT-1999; 99WO-US23089.

PR	29-NOV-1999;	99MO-US28214
PR	30-NOV-1999;	99MO-US28219
PR	30-NOV-1999;	99MO-US28409
PR	01-DEC-1999;	99MO-US28301
PR	01-DEC-1999;	99MO-US28634
PR	02-DEC-1999;	99MO-US28551
PR	02-DEC-1999;	99MO-US28564
PR	02-DEC-1999;	99MO-US28565
PR	16-DEC-1999;	99MO-US30911
PR	16-DEC-1999;	99MO-US30915
PR	20-DEC-1999;	99MO-US30929
PR	22-DEC-1999;	99MO-US30720
PR	30-DEC-1999;	99MO-US31243
PR	30-DEC-1999;	99MO-US31274
PR	05-JAN-2000;	2000MO-US00219
PR	06-JAN-2000;	2000MO-US00277
PR	11-FEB-2000;	2000MO-US00376
PR	18-FEB-2000;	2000MO-US04361
PR	18-FEB-2000;	2000MO-US04365
PR	18-FEB-2000;	2000MO-US04342
PR	22-FEB-2000;	2000MO-US04914
PR	24-FEB-2000;	2000MO-US04914
PR	01-MAR-2000;	2000MO-US05004
PR	02-MAR-2000;	2000MO-US05746
PR	02-MAR-2000;	2000MO-US05849
PR	02-MAR-2000;	2000MO-US05811
PR	15-MAR-2000;	2000MO-US06884
PR	15-MAR-2000;	2000MO-US07387
PR	21-MAR-2000;	2000MO-US07532
PR	17-MAY-2000;	2000MO-US08439
PR	30-MAY-2000;	2000MO-US13705
PR	30-MAY-2000;	2000MO-US14042
PR	30-MAY-2000;	2000MO-US14941
PR	28-JUN-2000;	2000MO-US15264
PR	11-AUG-2000;	2000MO-US20710
PR	23-AUG-2000;	2000MO-US22031
PR	23-AUG-2000;	2000MO-US23328
PR	10-NOV-2000;	2000MO-US30952
PR	10-NOV-2000;	2000MO-US30873
PR	01-DEC-2000;	2000MO-US32678
PR	28-FEB-2001;	2001MO-US05620
PR	01-MAR-2001;	2001MO-US06660
PR	25-MAY-2001;	2001MO-US10962
PR	20-JUN-2001;	2001MO-US17800
PR	22-JUN-2001;	2001MO-US20116
PR	09-JUN-2001;	2001MO-US21066
PR	29-JUN-2001;	2001MO-US21735
PR	28-DEC-2000;	2000MO-US74729
PR	28-FEB-2001;	2001MO-US76498
PR	09-MAR-2001;	2001MO-US07508
PR	12-MAR-2001;	2001MO-US08689
PR	05-APR-2001;	2001MO-US16744
PR	10-MAY-2001;	2001MO-US854208
PR	10-MAY-2001;	2001MO-US854280
PR	14-JUN-2001;	2001MO-US86216
PR	15-MAY-2001;	2001MO-US06628
PR	18-JUN-2001;	2001MO-US08879
PR	21-JUN-2001;	2001MO-US086827
PR	06-AUG-2001;	2001MO-US924419
PR	09-AUG-2001;	2001MO-US927796
PR	16-AUG-2001;	2001MO-US931836
PR	19-DEC-2001;	2001MO-US028072

BA	(GETH ) GENENTECH INC.
XX	
XX	Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W,
PI	Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S,
PI	Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
XX	WPI: 2003-333040/31.
DR	P-PSTDB; ABU66650.
XX	
PT	New secreted and transmembrane PRO nucleic acids, useful for gene
PT	therapy, in chromosome and gene mapping, as chromosome markers, in
PT	tissue typing, and in chromosome identification -
XX	
XX	Claim 2, Fig 16i: 66Op: English.
PS	
XX	The present invention relates to the isolation of novel human PRO
CC	polypeptides, and the polynucleotide sequences encoding them. The
CC	PRO polypeptides are secreted and transmembrane proteins. The PRO
CC	polypeptides are useful for detecting other PRO polypeptides, for
CC	linking bioactive molecules to cells expressing PRO polypeptides,
CC	for modulating biological activities of cells expressing PRO
CC	polypeptides, and for identifying agonists or antagonists.
CC	The PRO polypeptides are useful for stimulating the release of
CC	tumour necrosis factor (TNF)-alpha from human blood, for stimulating
CC	the proliferation or differentiation of chondrocytes, and detecting the
CC	presence of tumours. The polynucleotide sequences encoding PRO
CC	polypeptides are useful as hybridisation probes, in chromosome and
CC	gene mapping, in the generation of antisense RNA and DNA, in the
CC	preparation of PRO polypeptides, for generating transgenic animals or
CC	knockout animals, for the genetic analysis of individuals with genetic
CC	disorders, and in gene therapy. ACA03603-ACA03877 represent cDNAs
CC	encoding the human PRO polypeptides of the invention.
CC	Note: The sequence data for this patent was obtained in electronic
CC	format directly from the USPTO web site at
XX	seqdata.uspto.gov/psipdbidentity.html.
XX	
SQ	Sequence 1376 BP: 311 A; 384 C; 352 G; 329 T; 0 other:
	Query Match            18.1%; Score 1117; DB 25; Length 1376;
	Best Local Similarity 100.0%; Pred. No. 0;
	Matches 1117, Conservative 0; Mismatches 0; Indels 0; Gaps 0
OY	190 CCAGGGCGGAGCGCAGCTGCCCGGCGCTTGGGCGCTTGGGGCGCGCTCCACCGTCG 249
Db	1 CCAGGGCGGAGCGCAGCTGCCCGGCGCTTGGGCGCGCTTGGGGCGCGCTCCACCGTCG 60
OY	250 TTTTCCCCACGAGGCGGAGGCGTCGCCGGAATCATGCGCGGCGCTGAACTCGGGGTCT 309
Db	61 TTTTCCCACCGAAGCGGAGGCGTCGCCGGAATCATGCGCGGCGCTGAACTCGGGGTCT 120
OY	310 ATCCCACTGCATGAGGGGTTCTCTCTCGTGGGTCGCGGCGCTTGGCGGGGCAAGAAGT 369
Db	121 ATCCCACTGCATGAGGGGTTCTCTCTCGTGGGTCGCGGCGCTTGGCGGGGCAAGAAGT 180
OY	370 TTTCAGATTGCTGCCACGAGAAAACAATTAACAGTTCTATTAAGCTGGGGAACCCG 429
Db	181 TTTCAGATTGCTGCCACGAGAAAACAATTAACAGTTCTATTAAGCTGGGGAACCCG 240
OY	430 ACTCTGCTGGCAAAACCCTGTTAATCATGTCTATTTAAGAATATAAATGTTGTC 489
Db	.241 ACTCTGCTGGCAAAACCCTGTTAATCATGTCTATTTAAGAATATAAATGTTGTC 300
OY	490 ATCAAGTCTGGAGAAAGAAATAGTCTTTAACCTTTTAGCTGSCAGAGTCTGAGAAATCACTT 549
Db	301 ATCAAGTCTGGAGAAAGAAATAGTCTTTAACCTTTTAGCTGSCAGAGTCTGAGAAATCACTT 360
OY	550 GTCATATAGATCCAGAAAAAATATTGACTGTATGTTCAGGCCCATGTCCTTTGGGAGGTT 609
Db	361 GTCATATAGATCCAGAAAAAATATTGACTGTATGTTCAGGCCCATGTCCTTTGGGAGGTT 420
OY	610 CAGCTTACGCTTCGACATCGTTGTGCTTAACCTCAACAGAACTTTTATCTGGAGATGTC 669
Db	421 CAGCTTACGCTTCGACATCGTTGTGCTTAACCTCAACAGAACTTTTATCTGGAGATGTC 480

```
QY 670 AAAGCTATAGAGCATCGGTTTAGAGCTGCAAGTTTTCATCCCTCGCCCTGAGGAGATC 729
DB 481 AAAGCTATAGAGCATCGGTTTAGAGCTGCAAGTTTTCATCCCTCGCCCTGAGGAGATC 540
QY 730 GGTCCGGGTGAAGAGCTGCCCAAGAGATCACTCACTCACTCACTCACTCACTCACTCC 789
DB 541 GGTCCGGGTGAAGAGCTGCCCAAGAGATCACTCACTCACTCACTCACTCACTCACTCC 600
QY 790 ACCGTTGAGAGTGGAACTTCTGACAGCAATGGCACTGTGTCCCGGATCAAGATGCA 849
DB 601 ACCGTTGAGAGTGGAACTTCTGACAGCAATGGCACTGTGTCCCGGATCAAGATGCA 660
QY 850 GAAGAGTGAAGAGCTTAAACCTCCAGAGTTCCACCCAGAAATGCTCCGCTTC 909
DB 661 GAAGAGTGAAGAGCTTAAACCTCCAGAGTTCCACCCAGAAATGCTCCGCTTC 720
QY 910 AGCATTTGAAACCGCTCATCTATTAAGAGTCTGTGATCATGAGTCTGTGAGGAT 969
DB 721 AGCATTTGAAACCGCTCATCTATTAAGAGTCTGTGATCATGAGTCTGTGAGGAT 780
QY 970 GAAGGCTGCAACCTGATGTCTGCACTACCCAGAGGCTCCCTGAGATGAGCTC 1029
DB 781 GAAGGCTGCAACCTGATGTCTGCACTACCCAGAGGCTCCCTGAGATGAGCTC 840
QY 1030 ATGAGGTGAGAGTGTGCTGCTGCACTGCGGGCAGGCTCTCTCTCAACTTC 1089
DB 841 ATGAGGTGAGAGTGTGCTGCTGCACTGCGGGCAGGCTCTCTCTCAACTTC 900
QY 1090 AACCTCTCAACTGTGAGAGAGAGAGCGGGTTGAATACTACATCCCGGCTTCAAC 1149
DB 901 AACCTCTCAACTGTGAGAGAGAGAGCGGGTTGAATACTACATCCCGGCTTCAAC 960
QY 1150 ACCAAGCCGAGGTGTTCAAGCTGAGAGCAAGCAAGCCCTGGAACTGTGCGGAACTTC 1209
DB 961 ACCAAGCCGAGGTGTTCAAGCTGAGAGCAAGCAAGCCCTGGAACTGTGCGGAACTTC 1020
QY 1210 AACCTCTCTGCAAGAGCTGTGACCAAGATGCAAGGATCTCCGGCTGAC 1269
DB 1021 AACCTCTCTGCAAGAGCTGTGACCAAGATGCAAGGATCTCCGGCTGAC 1080
QY 1270 TTCCAAGTTTGGTCCACATCCACAAATGAAGA 1306
DB 1081 TTCCAAGTTTGGTCCACATCCACAAATGAAGA 1117

RESULT 8
ACAO4104
ID ACAO4104 standard; cDNA; 1376 BP.
AC ACAO4104;
XX
XX
XX 27-MAY-2003 (first entry)
DE Human cDNA encoding a secreted/transmembrane protein, SEQ ID 161.
XX
XX
XX Human; ss; gene: secreted protein; transmembrane protein; PRO;
XX inflammatory disease; organ failure; atherosclerosis; cardiac injury;
XX infertility; birth defects; premature aging; AIDS; biosensor;
XX acquired immunodeficiency syndrome; cancer; diabetic complication;
XX bioreactor; tumour.
XX
XX Homo sapiens.
XX
XX US2003032155-A1.
XX
XX 13-FEB-2003.
XX
XX 03-MAY-2002; 2002US-0137865.
XX
XX 31-MAR-1997; 97WO-US05230.
XX 12-JUN-1998; 98WO-US12456.
XX 14-JUL-1998; 98WO-US14552.
PR 28-AUG-1998; 98WO-US17888.
PR 10-SEP-1998; 98WO-US18824.
PR 14-SEP-1998; 98WO-US19093.
PR 14-SEP-1998; 98WO-US19094.
PR 14-SEP-1998; 98WO-US19177.
PR 16-SEP-1998; 98WO-US19330.
PR 17-SEP-1998; 98WO-US19437.
PR 07-OCT-1998; 98WO-US21141.
PR 29-OCT-1998; 98WO-US22991.
PR 29-OCT-1998; 98WO-US22992.
PR 20-NOV-1998; 98WO-US24855.
PR 01-DEC-1998; 98WO-US25108.
PR 05-JAN-1999; 99WO-US00106.
PR 08-MAR-1999; 99WO-US05028.
PR 10-MAR-1999; 99WO-US05190.
PR 20-APR-1999; 99WO-US08615.
PR 14-MAY-1999; 99WO-US10733.
PR 02-JUN-1999; 99WO-US12552.
PR 01-SEP-1999; 99WO-US20111.
PR 08-SEP-1999; 99WO-US20594.
PR 13-SEP-1999; 99WO-US20944.
PR 15-SEP-1999; 99WO-US21090.
PR 15-SEP-1999; 99WO-US21547.
PR 05-OCT-1999; 99WO-US23089.
PR 29-NOV-1999; 99WO-US28214.
PR 30-NOV-1999; 99WO-US28313.
PR 01-DEC-1999; 99WO-US28301.
PR 01-DEC-1999; 99WO-US28634.
PR 02-DEC-1999; 99WO-US28551.
PR 02-DEC-1999; 99WO-US28564.
PR 02-DEC-1999; 99WO-US28565.
PR 16-DEC-1999; 99WO-US30095.
PR 20-DEC-1999; 99WO-US30911.
PR 20-DEC-1999; 99WO-US30999.
PR 22-DEC-1999; 99WO-US30720.
PR 30-DEC-1999; 99WO-US31243.
PR 05-JAN-1999; 99WO-US31274.
PR 05-JAN-2000; 2000WO-US00219.
PR 06-JAN-2000; 2000WO-US00277.
PR 06-JAN-2000; 2000WO-US00376.
PR 11-FEB-2000; 2000WO-US00565.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US04914.
PR 24-FEB-2000; 2000WO-US05004.
PR 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US05746.
PR 10-MAR-2000; 2000WO-US05841.
PR 10-MAR-2000; 2000WO-US06319.
PR 15-MAR-2000; 2000WO-US06884.
PR 20-MAR-2000; 2000WO-US07377.
PR 21-MAR-2000; 2000WO-US07532.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 28-JUL-2000; 2000WO-US20710.
PR 11-AUG-2000; 2000WO-US22031.
PR 23-AUG-2000; 2000WO-US23522.
PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 10-NOV-2000; 2000WO-US30873.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000WO-US34956.
PR 28-FEB-2001; 2001WO-US06520.
PR 01-MAR-2001; 2001WO-US06666.
PR 25-MAY-2001; 2001WO-US17092.
PR 01-JUN-2001; 2001WO-US17800.
PR 20-JUN-2001; 2001WO-US19692.
PR 22-JUN-2001; 2001WO-US20116.
```







DB 1021 AACCTCTCTCTCAAGGCTGTACCAAGTGCACAGGATCTCTCGGCTGCAG 1080  
Oy 1270 TTCCAAGTTTGGTCCACATCCACAAATGAAGCA 1306  
DB 1081 TTCCAAGTTTGGTCCACATCCACAAATGAAGCA 1117

RESULT 9  
ABX89221  
ID ABX89221 standard; cDNA; 1376 BP.  
XX  
AC ABX89221;  
XX  
DT 13-MAY-2003 (first entry)  
XX  
DE DNA encoding novel secreted and transmembrane protein PRO5773.  
XX  
KW Human; PRO; hypertrophy of neonatal heart; angiogenesis; wound healing;  
KW cardiac insufficiency disorder; cancer; tumour; immune response;  
KW adrenal cortical capillary endothelial growth; c-fos induction;  
KW vascular endothelial growth factor inhibition; VEGF inhibition;  
KW endothelial cell growth inhibitor; T-lymphocytes stimulation;  
KW retinal neurons cell survival; rod photoreceptor cell survival;  
KW retinal disorder; retinitis pigmentosa; kidney disorder;  
KW mammalian kidney mesangial cell proliferation; Berger disease;  
KW dermatitis; herpeticiformis; Crohn's disease; chondrocyte proliferation;  
KW chondrocyte redifferentiation; sports injury; arthritis; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN US2003017563-A1.  
XX  
PD 23-JAN-2003.  
XX  
PF 07-MAY-2002; 2002US-0140808.  
XX  
XX 31-MAR-1997; 97WO-US05230.  
PR 12-JUN-1998; 98WO-US12456.  
PR 14-JUL-1998; 98WO-US14552.  
PR 28-AUG-1998; 98WO-US17888.  
PR 10-SEP-1998; 98WO-US18824.  
PR 14-SEP-1998; 98WO-US19093.  
PR 14-SEP-1998; 98WO-US19094.  
PR 14-SEP-1998; 98WO-US19177.  
PR 16-SEP-1998; 98WO-US19330.  
PR 17-SEP-1998; 98WO-US19437.  
PR 07-OCT-1998; 98WO-US21411.  
PR 29-OCT-1998; 98WO-US22991.  
PR 29-OCT-1998; 98WO-US22992.  
PR 20-NOV-1998; 98WO-US24855.  
PR 01-DEC-1998; 98WO-US25108.  
PR 05-JAN-1999; 99WO-US00106.  
PR 08-MAR-1999; 99WO-US05028.  
PR 10-MAR-1999; 99WO-US05190.  
PR 20-APR-1999; 99WO-US08615.  
PR 14-MAY-1999; 99WO-US10733.  
PR 02-JUN-1999; 99WO-US12252.  
PR 01-SEP-1999; 99WO-US20111.  
PR 08-SEP-1999; 99WO-US20594.  
PR 13-SEP-1999; 99WO-US20944.  
PR 15-SEP-1999; 99WO-US21090.  
PR 15-SEP-1999; 99WO-US21547.  
PR 05-OCT-1999; 99WO-US23089.  
PR 29-NOV-1999; 99WO-US28214.  
PR 30-NOV-1999; 99WO-US28313.  
PR 30-NOV-1999; 99WO-US28409.  
PR 01-DEC-1999; 99WO-US28301.  
PR 01-DEC-1999; 99WO-US28634.  
PR 02-DEC-1999; 99WO-US28551.  
PR 02-DEC-1999; 99WO-US28564.  
PR 02-DEC-1999; 99WO-US28565.  
PR 16-DEC-1999; 99WO-US30095.

PR 20-DEC-1999; 99WO-US30911.  
PR 20-DEC-1999; 99WO-US30999.  
PR 22-DEC-1999; 99WO-US30720.  
PR 30-DEC-1999; 99WO-US31243.  
PR 30-DEC-1999; 99WO-US31274.  
PR 05-JAN-2000; 2000WO-US00219.  
PR 06-JAN-2000; 2000WO-US00277.  
PR 11-FEB-2000; 2000WO-US03565.  
PR 18-FEB-2000; 2000WO-US04341.  
PR 18-FEB-2000; 2000WO-US04342.  
PR 22-FEB-2000; 2000WO-US04414.  
PR 24-FEB-2000; 2000WO-US04914.  
PR 24-FEB-2000; 2000WO-US05004.  
PR 01-MAR-2000; 2000WO-US05601.  
PR 02-MAR-2000; 2000WO-US05746.  
PR 02-MAR-2000; 2000WO-US05841.  
PR 10-MAR-2000; 2000WO-US06319.  
PR 15-MAR-2000; 2000WO-US06884.  
PR 20-MAR-2000; 2000WO-US07377.  
PR 21-MAR-2000; 2000WO-US07532.  
PR 30-MAR-2000; 2000WO-US08439.  
PR 17-MAY-2000; 2000WO-US13705.  
PR 22-MAY-2000; 2000WO-US14042.  
PR 30-MAY-2000; 2000WO-US14941.  
PR 02-JUN-2000; 2000WO-US15264.  
PR 28-JUL-2000; 2000WO-US20710.  
PR 11-AUG-2000; 2000WO-US22031.  
PR 23-AUG-2000; 2000WO-US23522.  
PR 24-AUG-2000; 2000WO-US23328.  
PR 08-NOV-2000; 2000WO-US30952.  
PR 10-NOV-2000; 2000WO-US30873.  
PR 01-DEC-2000; 2000WO-US32678.  
PR 20-DEC-2000; 2000WO-US34956.  
PR 28-FEB-2001; 2001WO-US06520.  
PR 01-MAR-2001; 2001WO-US06666.  
PR 25-MAY-2001; 2001WO-US17092.  
PR 01-JUN-2001; 2001WO-US17800.  
PR 20-JUN-2001; 2001WO-US19692.  
PR 22-JUN-2001; 2001WO-US20116.  
PR 29-JUN-2001; 2001WO-US21066.  
PR 09-JUL-2001; 2001WO-US21735.  
PR 20-DEC-2000; 2000US-074259.  
PR 28-FEB-2001; 2001US-0796498.  
PR 09-MAR-2001; 2001US-0803706.  
PR 14-MAR-2001; 2001US-0806889.  
PR 22-MAR-2001; 2001US-0816744.  
PR 05-APR-2001; 2001US-0826366.  
PR 10-MAY-2001; 2001US-0854280.  
PR 18-MAY-2001; 2001US-0860216.  
PR 25-MAY-2001; 2001US-0866028.  
PR 01-JUN-2001; 2001US-0876034.  
PR 05-JUN-2001; 2001US-0872035.  
PR 14-JUN-2001; 2001US-0882636.  
PR 19-JUN-2001; 2001US-0885342.  
PR 21-JUN-2001; 2001US-0887879.  
PR 18-JUL-2001; 2001US-0908827.  
PR 06-AUG-2001; 2001US-0924419.  
PR 09-AUG-2001; 2001US-0927796.  
PR 16-AUG-2001; 2001US-0931836.  
PR 19-DEC-2001; 2001US-0028072.  
XX  
PA (GENTH ) GENENTECH INC.  
XX  
XX Baker KP, Beresini M, Deforge L, Desnoyers L, Pilyaroff B, Gao W;  
PI Gertsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;  
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WT, Zhang Z;  
XX WPI; 2003-148238/14.  
DR P-PsDB; ABUS9731.  
XX



PR 13-DEC-2000; 2000US-255281P.  
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX Schlegel R, Endege WO, Monahan JE;  
XX WPI; 2001-662795/76.  
XX  
XX Novel isolated nucleic acid molecule associated with cancerous state of  
PT prostate cells and correlating with presence of prostate cancer, useful  
PT for detecting presence of prostate cancer; stage of prostate cancer -  
XX  
XX Claim 1; Page 6580; 11750pp; English.  
XX  
XX The invention relates to an isolated nucleic acid molecule (I) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement. (I) is useful for:  
CC (a) assessing whether a patient is afflicted with prostate cancer;  
CC (b) monitoring the progression of prostate cancer in a patient;  
CC (c) assessing the efficacy of a test compound to inhibit prostate  
CC cancer in a patient;  
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer  
CC in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound;  
CC (g) determining whether prostate cancer has metastasized in a patient;  
CC (h) assessing the aggressiveness or indolence of prostate cancer in a  
CC patient;  
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.  
XX  
XX Sequence 1031 BP; 272 A; 255 C; 212 G; 263 T; 29 other;  
XX  
XX  
XX Query Match 10.9%; Score 669; DB 23; Length 1031;  
XX Best Local Similarity 100.0%; Pred. NO. 1.9e-293;  
XX Matches 669; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
XX 4499 CTTTTCCTTTTCTTTTATGAGACAGAGATCTGTTGCTAGGCTGAGTGC 4558  
XX 30 CTTTTCCTTTTCTTTTATGAGACAGAGATCTGTTGCTAGGCTGAGTGC 89  
XX  
XX 4559 AGTGGCGCAATCTGGCTCACTGCAACCTCTGCTCTGGGCTCAAGCAATTTCTCCACC 4618  
XX 90 AGTGGCGCAATCTGGCTCACTGCAACCTCTGCTCTGGGCTCAAGCAATTTCTCCACC 149  
XX  
XX 4619 TTAGCTCTCCCAATAGCTGGAGTCACTGGACACAAACACCAATGCTTAATTTTGTAT 4678  
XX 150 TTAGCTCTCCCAATAGCTGGAGTCACTGGACACAAACACCAATGCTTAATTTTGTAT 209  
XX  
XX 4679 TTTTGTAGAGACAGAGGTTTCAACATGTTGCCAGGCTGATCTCAACTCTGAGGCTCAA 4738  
XX 210 TTTTGTAGAGACAGAGGTTTCAACATGTTGCCAGGCTGATCTCAACTCTGAGGCTCAA 269  
XX  
XX 4739 GCAATCTCTGCTGCTGGCTCTCCCAAGTGTGGATTAAGATGTGAGCCACCGCATCC 4798  
XX 270 GCAATCTCTGCTGCTGGCTCTCCCAAGTGTGGATTAAGATGTGAGCCACCGCATCC 329  
XX  
XX 4799 ACCCCACACCCCTCATTTATACCAATTAACCTGCCCCAGTAACTGACTTTTCTTCTTC 4858  
XX 330 ACCCCACACCCCTCATTTATACCAATTAACCTGCCCCAGTAACTGACTTTTCTTCTTC 389  
XX  
XX 4859 ACCCCCTGCTGATCTGGAAGAGAGAGATTAATGATAGCTTGTGAGACAGATCCCAAG 4918  
XX 390 ACCCCCTGCTGATCTGGAAGAGAGAGATTAATGATAGCTTGTGAGACAGATCCCAAG 449  
XX  
XX 4919 TTCAATATTTCTGCGGCAAAAATCTCTTCAAAAATAATGATCTTCAATTTTATTTCAAT 4978  
XX 450 TTCAATATTTCTGCGGCAAAAATCTCTTCAAAAATAATGATCTTCAATTTTATTTCAAT 509  
XX  
XX 4979 GAATTCACCTTGGAAATGACCGCCCTCACTGTTTCACTGATGATTAAGAAAGAAATTT 5038  
XX 510 GAATTCACCTTGGAAATGACCGCCCTCACTGTTTCACTGATGATTAAGAAAGAAATTT 569  
XX  
XX 5039 TATAGTCTCTAAATGCGGTGACTGCAAGACCTTTTGAACACTTTCAGAGAGATGAGAT 5098

DB 570 TATAGTCTCTAAATGCGGTGACTGCAAGACCTTGAACACTTTCAGAGAGATGAGAT 629  
XX  
XX 5099 ATTTAAGTCAGCCCTTGGCCGCTGCTAATGACACTTCCCTTGAAGTGTGTTCT 5158  
XX 630 ATTTAAGTCAGCCCTTGGCCGCTGCTAATGACACTTCCCTTGAAGTGTGTTCT 689  
XX  
XX 5159 GCCCAGTGA 5167  
XX 690 GCCCAGTGA 698  
XX  
XX  
XX RESULT 11  
XX ID AAA78075 standard; cDNA; 544 BP.  
XX  
XX AAA78075;  
XX  
XX 14-NOV-2000 (first entry)  
XX  
XX cDNA encoding human colon tumour polypeptide, SEQ ID NO:362.  
XX  
XX Human colon tumour polypeptide; tumour antigen; cancer; vaccine;  
XX immunotherapy; diagnosis; progression; ss.  
XX  
XX Homo sapiens.  
XX  
XX WO200037643-A2.  
XX  
XX 29-JUN-2000.  
XX  
XX 23-DEC-1999; 99WO-US30909.  
XX  
XX 23-DEC-1998; 98US-0221298.  
XX 02-JUL-1999; 99US-0347496.  
XX 22-SEP-1999; 99US-0401064.  
XX 19-NOV-1999; 99US-0444242.  
XX 02-DEC-1999; 99US-0454150.  
XX  
XX (CORI-) CORIXA CORP.  
XX  
XX Xu J, Lodes MJ, Secretat H, Benson DR, Meagher MJ, Stolk J;  
XX Wang T, Yugu J;  
XX  
XX WPI; 2000-442671/38.  
XX  
XX New colon tumor polypeptides used to inhibit the development of cancer,  
XX especially colon cancer, and for diagnosing and monitoring the  
XX progression of the cancer -  
XX  
XX Claim 1; Page 196; 229pp; English.  
XX  
XX Sequences AAA77722-A78199 represent 478 cDNAs encoding proteins or  
XX portions of proteins which are associated with human colon tumours.  
XX The invention also specifically discloses 8 human colon tumour proteins  
XX (AAH1897-B11904). The nucleic acids, the polypeptides they encode, and  
XX antigen presenting cells (APCs; preferably dendritic cells) expressing  
XX such polypeptides may be used in vaccines that target tumour cells,  
XX especially colon tumour cells, thereby inhibiting the development of  
XX cancer. T-cells specific for the polypeptide expressed by the APC are  
XX used to remove tumour cells from biological samples, especially blood or  
XX fractions thereof. The sample or the isolated T-cells specific for the  
XX polypeptide can then be used to inhibit cancer development. CD4+ and/or  
XX CD8+ T-cells from a patient may be incubated with a polypeptide or  
XX nucleic acid of the invention, or an APC expressing such a polypeptide,  
XX to cause the proliferation of specific T-cells. The T-cells can be  
XX cloned and then administered back to the patient to inhibit cancer  
XX development. Nucleic acids encoding the polypeptides and antibodies  
XX against the polypeptides may be used to determine the expression level  
XX of a tumour protein of the invention, and therefore to determine whether  
XX cancer cells are present. Such diagnostic methods may also be used to  
XX monitor the progression of a cancer by repeating the processes at time  
XX intervals, and comparing the current result to previous results. The



Qy	3428	TGTTTGAGCAGCATTTGACACATATCTGTTTGATTAAGACATTCCTGATTCCTAGTCG	3487
Db	361	TGTTTGAGCAGCATTTGACACATATCTGTTTGATTAAGACATTCCTGATTCCTAGTCG	420
Qy	3488	GTTCGTGGTTATCCCATTTGTGGAATTCATCTTGATTCATATGCTATAGTCTTAGCA	3547
Db	421	GTTCGTGGTTATCCCATTTGTGGAATTCATCTTGATTCATATGCTATAGTCTTAGCA	480
Qy	3548	ATAAGAGAAATTTCTCTCAAGTTTCCATGTGGGTTCTCTAGCTGACCAATACCTTTCAC	3607
Db	481	ATAAGAGAAATTTCTCTCAAGTTTCCATGTGGGTTCTCTAGCTGACCAATACCTTTCAC	540
Qy	3608	ATTT 3611	
Db	541	ATTT 544	
RESULT 13			
ID	AB232999	standard; cDNA; 544 BP.	
XX	AB232999;		
XX	30-JAN-2003	(first entry)	
DE	Human colon tumour cDNA clone 25928 SEQ ID NO:362.		
XX	Human; colon cancer; colon tumour; immunotherapy; diagnosis; cancer;		
KW	tumour; immune response; immunostimulant; cytostatic; vaccine;		
XX	gene; ss.		
OS	Homo sapiens.		
XX	MO200283070-A2.		
PD	24-OCT-2002.		
XX	09-APR-2002; 2002MO-US11475.		
PF	10-APR-2001; 2001US-0833263.		
XX	PR 03-AUG-2001; 2001US-0922217.		
PR	19-DEC-2001; 2001US-0025380.		
XX	(CORI-) CORIXA CORP.		
PA			
XX	Xu J, Lodes MJ, Secrist H, Benson DR, Meagher MJ, Stolk JA;		
PI	Wang T, Jiang Y, Smith CL, King GE, Wang A, Clapper JD;		
PI	Skelky YAM, Fanger GR, Vedvick TS, Carter D;		
XX	WPI; 2003-067548/06.		
DR			
XX			
PT	New polynucleotide, useful for the preparation of a composition for		
PT	stimulating an immune response against, or treating, cancer -		
XX			
PS	Example 1; Page 236-237; 537pp; English.		
XX			
CC	The present invention describes compounds (I) for the immunotherapy and		
CC	diagnosis of colon cancer. Also described: (1) a method for detecting		
CC	the presence of cancer in a patient; (2) a method for stimulating and/or		
CC	expanding T cells specific for a tumour protein; (3) an isolated T cell		
CC	population comprising T cells prepared by the method of (2); (4) a method		
CC	for stimulating an immune response in a patient; (5) a method for		
CC	treating cancer in a patient; and (6) a method for inhibiting the		
CC	development of cancer in a patient. (I) have immunostimulant and		
CC	cytostatic activities and can be used in vaccines. AB232646 to AB233725		
CC	and ABP55343 to ABP55351 represent human colon cancer/tumour related		
CC	sequences used in the exemplification of the present invention.		
XX			
XX	Sequence 544 BP; 138 A; 123 C; 113 G; 170 T; 0 other;		
Qy	Query Match	8.0%; Score 493; DB 25; Length 544;	
XX	Best Local Similarity	99.8%; Pred. No. 1.9e-213;	
XX	Matches 543; Conservative 0; Mismatches 1; Indels 0; Gaps 0;		

Oy	3068	CCTAGTACCTACCTACGATGAGGTTTCACAGAACCCCTGGATTCCAGAGTGTAAACAGAGGCT	3127
Db	1	CTTAGTACCTACCTACGATGAGGTTTCACAGAACCCCTGGATTCCAGAGTGTAAACAGAGGCT	60
Oy	3128	TGCCCTCTTTCAGACCAACAGTTCCAAATTCACAAGAGGCTTACCTGAGGTCCTTACTCTCAC	3187
Db	61	TGCCCTCTTTCAGACCAACAGTTCCAAATTCACAAGAGGCTTACCTGAGGTCCTTACTCTCAC	120
Oy	3188	TGGGGTCCCCAGGATGAAAAAGAAATGTCCTTTTATATATATATATATTTGGTGTCCT	3247
Db	121	TGGGGTCCCCAGGATGAAAAAGAAATGTCCTTTTATATATATATATATTTGGTGTCCT	180
Oy	3248	GTGTTATTTTAAAGATCAAAATGTATTAACCACTTACCTTTTCACTGACTTACTAGTAA	3307
Db	181	GTGTTATTTTAAAGATCAAAATGTATTAACCACTTACCTTTTCACTGACTTACTAGTAA	240
Oy	3308	CTCAATCTAACTGGTTTGGATGCTCGGAGTTGTGACTTCTACTGACCGCTAGATTAACGCTG	3367
Db	241	CTCAATCTAACTGGTTTGGATGCTCGGAGTTGTGACTTCTACTGACCGCTAGATTAACGCTG	300
Oy	3368	TGCCTGTCCCCAGAGTGTGGGAATATTTTACAATCTGTCCAACCAAAAAAGATGTGCTG	3427
Db	301	TGCCTGTCCCCAGAGTGTGGGAATATTTTACAATCTGTCCAACCAAAAAAGATGTGCTG	360
Oy	3428	TGTTTGACGACGATTTGACACATATCTGCTTTTGTATTAAGAGACTTCTGATTTCTTACGCTG	3487
Db	361	TGTTTGACGACGATTTGACACATATCTGCTTTTGTATTAAGAGACTTCTGATTTCTTACGCTG	420
Oy	3488	GTTCGTGTGTTATCCCATTTGTGAAAAATTCATCTTGATCCCATTTGCTTATATGTCCTTACGA	3547
Db	421	GTTCGTGTGTTATCCCATTTGTGAAAAATTCATCTTGATCCCATTTGCTTATATGTCCTTACGA	480
Oy	3548	ATAGAGAGAAATTCCTCAAGTTCCATGTCGGTTCCTTACGTCGACGAATACTTTGAC	3607
Db	481	ATAGAGAGAAATTCCTCAAGTTCCATGTCGGTTCCTTACGTCGACGAATACTTTGAC	540
Oy	3608	ATTTT 3611	
Db	541	ATTTT 544	
RESULT 14			
ABT21661/C			
ID	ABT21661	standard; DNA; 660 BP.	
AC	ABT21661;		
XX			
DT	16-APR-2003	(first entry)	
XX			
DE	Breast cancer marker gene SEQ ID No 34.		
XX			
KM	Cytosratic; vaccine; breast cancer marker gene; breast mass; immunogen;		
KW	chemotherapy; tumour burden; balt protein; two-hybrid; three-hybrid;		
KW	surrogate marker gene; pharmacodynamic marker gene; transgenic animal;		
XX	human; ds.		
OS	Homo sapiens.		
XX			
XX	WO200285298-A2.		
XX			
PD	31-OCT-2002.		
XX			
PF	19-APR-2002; 2002WO-US12612.		
XX			
PR	20-APR-2001; 2001US-285163P.		
XX			
PA	(MILL-) MILLENNIUM PHARM INC.		
XX			
PI	Lillie J, Palermo A, Wang Y, Steinmann K, Elias J, Mertens M;		
XX			
DR	WPI; 2003-093053/08.		
XX			

PT Novel isolated polypeptide encoded by breast cancer marker gene, useful  
PT for diagnosing, staging, monitoring, prognosing and treating diseases  
PT associated with breast cancer  
PS Disclosure, Page 101; 725pp; English.  
XX  
XX  
CC The invention relates to an isolated polypeptide encoded by a breast  
CC cancer marker gene comprising any of 1417 21-805 nucleotide sequences,  
CC given in the specification. The methods of the invention are useful for  
CC diagnosing patients having an identified breast mass or symptoms  
CC associated with breast cancer, to diagnose breast cancer or its  
CC precursors, and for monitoring the efficacy of treatment of a breast  
CC cancer patient (e.g. efficacy of chemotherapy). The methods are also  
CC useful for evaluating a patient before, after or during therapy, to  
CC evaluate the reduction in a tumour burden. The breast cancer marker gene  
CC proteins are useful as immunogens for raising antibodies, by immunising a  
CC mammal with a breast cancer marker protein. The marker proteins are  
CC useful as bait proteins in a two-hybrid or three-hybrid assay, to  
CC identify other proteins which bind to or interact with the marker  
CC proteins. The breast cancer marker genes are useful as surrogate marker  
CC genes for one or more disorders, disease states or conditions leading to  
CC disease states, in particular, breast cancers. The breast cancer marker  
CC genes are useful as pharmacodynamic marker genes. An antibody which  
CC selectively binds to a protein of a breast cancer marker gene is useful  
CC for treating cancers, particularly breast cancers. The host cell of the  
CC invention is useful for producing non-human transgenic animals. This  
CC polynucleotide sequence represents one of the breast cancer marker genes  
CC of the invention.  
XX  
XX  
SQ Sequence 660 BP; 162 A; 159 C; 182 G; 154 T; 3 other;

Query Match 7.8%; Score 486; DB 25; Length 660;  
Best Local Similarity 100.0%; Pred. No. 2.9e-210;  
Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5421 CCACGCGGCGCATTTTCTTAACTGAGGGCTCAAACTCCTGGACAAGTTGCTGCTCT 5480  
DB 532 CCACGCGGCGCATTTTCTTAACTGAGGGCTCAAACTCCTGGACAAGTTGCTGCTCT 473  
QY 5481 GAGACCGATTTCTCGAGCTGAGCTCAGTAAAGGGGCCAGCCTGAGGAAGCCCTGGC 5540  
DB 472 GAGACCGATTTCTCGAGCTGAGCTCAGTAAAGGGGCCAGCCTGAGGAAGCCCTGGC 413  
QY 5541 TCTTTTCTTAAAGCCAGGCCCACTTACATTAACATTTTCAAGGGTCACTGGAAACAGT 5600  
DB 412 TCTTTTCTTAAAGCCAGGCCCACTTACATTAACATTTTCAAGGGTCACTGGAAACAGT 353  
QY 5601 GAAGTGCATTTTGTGAAGCTCTGATGCCAGCCCACTGCTCATCTCAAGTGTCTGCC 5660  
DB 352 GAAGTGCATTTTGTGAAGCTCTGATGCCAGCCCACTGCTCATCTCAAGTGTCTGCC 293  
QY 5661 ATCCCTACGAGGAAGGCCAGCGATGAGAGCTGCTCAATGCTGTGTCATTCACAA 5720  
DB 292 ATCCCTACGAGGAAGGCCAGCGATGAGAGCTGCTCAATGCTGTGTCATTCACAA 233  
QY 5721 GAAGGGAAGAGTCTCAAGGAAGTCAACTGAGGACAAGCAAGCCCAAGCATGGCC 5780  
DB 232 GAAGGGAAGAGTCTCAAGGAAGTCAACTGAGGACAAGCAAGCCCAAGCATGGCC 173  
QY 5781 TTGGTAAAGTTTGCAGACTGCTGTGTGTGATCTGCAAGTCTTCACTGGAATTAATTTA 5840  
DB 172 TTGGTAAAGTTTGCAGACTGCTGTGTGTGATCTGCAAGTCTTCACTGGAATTAATTTA 113  
QY 5841 TTCAATTCAGATCTTTTATGAGTGGCATTTTATTCATTTCTGCTTTAAATTAACAAA 5900  
DB 112 TTCAATTCAGATCTTTTATGAGTGGCATTTTATTCATTTCTGCTTTAAATTAACAAA 53  
QY 5901 TGTACC 5906  
DB 52 TGTACC 47

RESULT 15

AAS64483  
ID AAS64483 standard; cDNA; 487 BP.  
XX  
XX AAS64483;  
AC  
XX  
DT 13-FEB-2002 (first entry)  
DE  
XX DNA encoding novel human diagnostic protein #287.  
XX  
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.  
XX  
XX Homo sapiens.  
XX  
XX WO200175067-A2.  
XX  
XX 11-OCT-2001.  
XX  
XX 30-MAR-2001; 2001WO-US08631.  
XX  
XX 31-MAR-2000; 2000US-0540217.  
XX 23-AUG-2000; 2000US-0649167.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
XX Dormanac RT, Liu C, Tang YT;  
XX  
XX WPI; 2001-639362/73.  
XX P-PsDB; ABG00296.  
XX  
XX  
XX New isolated polynucleotide and encoded polypeptides, useful in  
XX diagnostics, forensics, gene mapping, identification of mutations  
XX PT responsible for genetic disorders or other traits and to assess  
XX PT biodiversity  
XX  
XX Claim 1, SEQ ID No 287; 103pp; English.  
PS  
XX  
XX The invention relates to isolated polynucleotide (I) and  
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,  
XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
XX CC and gene mapping, and in recombinant production of (II). The  
XX CC polynucleotides are also used in diagnostics as expressed sequence tags  
XX CC for identifying expressed genes. (I) is useful in gene therapy techniques  
XX CC to restore normal activity of (II) or to treat disease states involving  
XX CC (II). (II) is useful for generating antibodies against it, detecting or  
XX CC quantitating a polypeptide in tissue, as molecular weight markers and as  
XX CC a food supplement. (II) and its binding partners are useful in medical  
XX CC imaging of sites expressing (II). (I) and (II) are useful for treating  
XX CC disorders involving aberrant protein expression or biological activity.  
XX CC The polypeptide and polynucleotide sequences have applications in  
XX CC diagnostics, forensics, gene mapping, identification of mutations  
XX CC responsible for genetic disorders or other traits to assess biodiversity  
XX CC and to produce other types of data and products dependent on DNA and  
XX CC amino acid sequences. AAS64197-AAS94564 represent novel human  
XX CC diagnostic coding sequences of the invention.  
XX CC Note: The sequence data for this patent did not appear in the printed  
XX CC specification, but was obtained in electronic format directly from WIPO  
XX CC at ftp.wipo.int/pub/published\_pct\_sequences.

SQ Sequence 487 BP; 131 A; 145 C; 120 G; 91 T; 0 other;

Query Match 7.8%; Score 482; DB 23; Length 487;  
Best Local Similarity 100.0%; Pred. No. 1.9e-208;  
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2312 GAGCTTACTGCTGTGCTGCGCTCGGCTCATATTTGCTGTGTGAAAAAAGAAAAA 2371  
DB 6 GAGCTTACTGCTGTGCTGCGCTCGGCTCATATTTGCTGTGTGAAAAAAGAAAAA 65  
QY 2372 AGACAACAAGAGGCGCCGCTGAGGATCTACAATGGCAACATCACTAGATGCGCA 2431  
DB 66 AGACAACAAGAGGCGCCGCTGAGGATCTACAATGGCAACATCACTAGATGCGCA 125

OY	2432	GGCAGCCAAAAAGTTTCAGAAAGGGCGAAAAGACAATGACCTCCCATGTGTATGACGCA	2431
Db	126	GGCAGCCAAAAAGTTTCAGAAAGGGCGAAAAGACAATGACCTCCCATGTGTATGACGCA	185
OY	2492	TCGAGGACCACTGGTATATGAGGCACTTGCTACAGAGATTCCAGCGGCTCTTCTCGACGC	2551
Db	186	TCGAGGACCACTGGTATATGAGGCACTTGCTACAGAGATTCCAGCGGCTCTTCTCGACGC	245
OY	2552	CAGAGGTGACACCTTACCGGCGCTTCCAGGGGACCATGGGGGTCTGTCTCTCCCTCCAC	2611
Db	246	CAGAGGTGACACCTTACCGGCGCTTCCAGGGGACCATGGGGGTCTGTCTCTCTCCCTCCAC	305
OY	2612	CCACCATATGCTCCAGGGGCCCAACTGTGAAAAGTTGGGCACTGAGGAGCACCTCTGGCT	2671
Db	306	CCACCATATGCTCCAGGGGCCCAACTGTGAAAAGTTGGGCACTGAGGAGCACCTCTCTGGCT	365
OY	2672	CCCTCTCTGAGTCTTGAGAGTGAAACCGTACACCTTCTCCCATCCCAACATGGGGATGTAA	2731
Db	366	CCCTCTCTGAGTCTTGAGAGTGAAACCGTACACCTTCTCCCATCCCAACATGGGGATGTAA	425
OY	2732	GCAGCAAGGACACAGACATTTCCCTTACTGAAACCTCAGGAGCCCATGGAGCCACAGCAAT	2791
Db	426	GCAGCAAGGACACAGACATTTCCCTTACTGAAACCTCAGGAGCCCATGGAGCCACAGCAAT	485
OY	2792	AA	2793
Db	486	AA	487

Search completed: February 20, 2004, 04:34:29  
Job time : 1489 BECS

**This Page Blank (uspio)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2004, 02:41:05 ; Search time 21696 Seconds

(without alignment)  
11620.853 Million cell updates/sec

Title: US-09-899-569A-3

Perfect score: 6163

Sequence: 1 ccaacgcgcgaatggggagtc.....agtcgcgcaaaaaaaaaa 6163

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 2888711 seqs, 2045481386 residues

Word size : 0

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: listing first 45 summaries

Database : GenEmbl:

1: gb\_ba:\*

2: gb\_hcg:\*

3: gb\_in:\*

4: gb\_om:\*

5: gb\_ov:\*

6: gb\_pac:\*

7: gb\_ph:\*

8: gb\_pl:\*

9: gb\_pr:\*

10: gb\_ro:\*

11: gb\_scs:\*

12: gb\_by:\*

13: gb\_un:\*

14: gb\_vl:\*

15: em\_ba:\*

16: em\_fun:\*

17: em\_hum:\*

18: em\_in:\*

19: em\_mu:\*

20: em\_om:\*

21: em\_or:\*

22: em\_ov:\*

23: em\_pat:\*

24: em\_ph:\*

25: em\_pl:\*

26: em\_ro:\*

27: em\_scs:\*

28: em\_un:\*

29: em\_vl:\*

30: em\_hcg\_hum:\*

31: em\_hcg\_inv:\*

32: em\_hcg\_other:\*

33: em\_hcg\_nub:\*

34: em\_hcg\_pln:\*

35: em\_hcg\_rtd:\*

36: em\_hcg\_mam:\*

37: em\_hcg\_vrt:\*

38: em\_hcg\_hum:\*

39: em\_hcg\_mus:\*

40: em\_hcg\_mus:\*

41: em\_hcg\_other:\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	6163	100.0	6163	6	AX353605 Sequence
2	5963	96.8	5963	6	AY026461 Homo sapi
3	5395	87.5	5897	6	AX353603 Sequence
4	4996	81.1	5978	6	AF468010 Homo sapi
5	4786	77.7	5573	6	BD160235 Primer fo
6	4786	77.7	5573	6	AK023834 Homo sapi
7	3268	53.0	163115	9	AC105902 Homo sapi
8	3268	53.0	200956	9	AC104165 Homo sapi
9	3004	48.7	3224	9	AK026622 Homo sapi
10	1703	27.6	2113	9	AK026028 Homo sapi
11	1545	25.1	2209	6	AR243778 Sequence
12	1177	19.1	1241	9	AK026187 Homo sapi
13	1159	18.8	1410	9	BC021099 Homo sapi
14	1117	18.1	1376	6	AX464028 Sequence
15	909	14.7	976	9	AK026329 Homo sapi
16	493	8.0	544	6	AX192795 Sequence
17	477	7.7	541	6	AX261684 Sequence
18	453	7.4	535	6	AX260742 Sequence
19	435	7.1	636	6	AX261015 Sequence
20	434	7.0	3301	6	AX576112 Sequence
21	404	6.6	659	6	BD150371 Primer fo
22	383	6.2	561	6	BD155313 Primer fo
23	364	5.9	173977	9	AC010170 Homo sapi
24	279	4.5	400	6	AX203285 Sequence
25	174	2.8	174	6	AR251796 Sequence
26	140	2.3	262	6	AX261252 Sequence
27	68	1.1	104531	9	HS0303A1 Human DNA
28	68	1.1	153241	9	AC023510 Homo sapi
29	68	1.1	162589	2	AC015717 Homo sapi
30	68	1.1	173585	9	AC090527 Homo sapi
31	68	1.1	218724	2	AC025889 Homo sapi
32	67	1.1	73775	9	AP003901 Homo sapi
33	67	1.1	107469	9	HS294 Human DNA
34	67	1.1	119182	9	HSB62L20 Homo sapi
35	67	1.1	162364	9	AP005213 Homo sapi
36	67	1.1	190690	2	AL627234 Homo sapi
37	67	1.1	340000	9	HS21C003 Homo sapi
38	64	1.0	155304	9	AC055845 Homo sapi
39	64	1.0	174187	2	AP002792 Homo sapi
40	64	1.0	215441	2	AP002800 Homo sapi
41	64	1.0	215647	2	AC068591 Homo sapi
42	63	1.0	724	6	AX417709 Sequence
43	63	1.0	106605	9	AL353691 Human DNA
44	63	1.0	132290	9	AL627402 Human DNA
45	63	1.0	144735	2	AL161458 Homo sapi

## ALIGNMENTS

RESULT 1	AX353605	6163 bp	DNA	linear	PAT 06-FEB-2002
LOCUS	AX353605				
DEFINITION	Sequence 3 from Patent WO0204508.				
ACCESSION	AX353605				
VERSION	AX353605.1	GI:18618678			
KEYWORDS					
SOURCE					
ORGANISM	Homo sapiens (human)				
REFERENCE					
AUTHORS	1				
TITLE	Schwefler,N., Scherl-Mostagier,M., Sommergruber,W. and Abseher,R.				
	Tumour-associated antigen (b345), characterised by an amino acid				
	sequence as in seq. Id. No. 4				

Pred. No. is the number of results predicted by chance to have a



1621 TCCTGGAAGCTGCTGTGTCCTGAGACAGGCTCAGCTGTGTGTGTGTCGACCCAGAG 1680  
1621 TCCTGGAAGCTGCTGTGTCCTGAGACAGGCTCAGCTGTGTGTGTGTCGACCCAGAG 1680  
1681 CTGACGACGACATACACAGAGAGCCCTGACACACAGCTTACCTGTGTGTCGAGT 1740  
1681 CTGACGACGACATACACAGAGAGCCCTGACACACAGCTTACCTGTGTGTCGAGT 1740  
1741 GCCATACCAGCAGCAGACCTGTACTTCTGAGCTCTCTGAGGAGGAGCTTATCAAGAG 1800  
1741 GCCATACCAGCAGCAGACCTGTACTTCTGAGCTCTCTGAGGAGGAGCTTATCAAGAG 1800  
1801 ATCCAGGTGAAGCAGACATCTCGGTGACCTTCCGACCTTTGCCCCAGCTTCCAAACA 1860  
1801 ATCCAGGTGAAGCAGACATCTCGGTGACCTTCCGACCTTTGCCCCAGCTTCCAAACA 1860  
1861 GAGGCTCCAGCAGGAGCTGTGACGGTGTCTTATACCTTATTTTAAAGAGAGAGCGTT 1920  
1861 GAGGCTCCAGCAGGAGCTGTGACGGTGTCTTATACCTTATTTTAAAGAGAGAGCGTT 1920  
1921 TTCAAGGTGACCCCTGACACAAAGAGGTCTACTGAGAGACCCCACTGGAGACCG 1980  
1921 TTCAAGGTGACCCCTGACACAAAGAGGTCTACTGAGAGACCCCACTGGAGACCG 1980  
1981 GGCCTGCCATCCCTCACCCTGTGTCTGTGAACATCAGCGTCCAGAGACAGGTGAGC 2040  
1981 GGCCTGCCATCCCTCACCCTGTGTCTGTGAACATCAGCGTCCAGAGACAGGTGAGC 2040  
2041 TGCCGTGACCTTTTAAAGAGAGCGGCGGTGTGTGCAAGAGGCGCCATTTCAATG 2100  
2041 TGCCGTGACCTTTTAAAGAGAGCGGCGGTGTGTGCAAGAGGCGCCATTTCAATG 2100  
2101 ATCATCCAGAGACAGCGGACCGGCGGTGTGAGATCTTCAGCGCTGAGAGAGTGTGCTC 2160  
2101 ATCATCCAGAGACAGCGGACCGGCGGTGTGAGATCTTCAGCGCTGAGAGAGTGTGCTC 2160  
2161 CCCAAGCCAGGCTTCACCATCAGAGCTTCTGTGATCAATCTCTTACCTGACGCCCAAG 2220  
2161 CCCAAGCCAGGCTTCACCATCAGAGCTTCTGTGATCAATCTCTTACCTGACGCCCAAG 2220  
2221 AGCGGCAAGACGCTGACCTGTCTTCTGTGATCAATCTCTTACCTGACGCCCAAG 2280  
2221 AGCGGCAAGACGCTGACCTGTCTTCTGTGATCAATCTCTTACCTGACGCCCAAG 2280  
2281 ACTGTCATCTCTGACGAGCGGTGTGAGAGTCTTACCTGACGCCCAAG 2340  
2281 ACTGTCATCTCTGACGAGCGGTGTGAGAGTCTTACCTGACGCCCAAG 2340  
2341 ATCATTTGCTGTGTGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGTGTGTGATTC 2400  
2341 ATCATTTGCTGTGTGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGTGTGTGATTC 2400  
2401 TACAAATGGCAATCAATCTGAGATGCGGAGGAGCCAAAAAGTTTCAAGAAAGGCGA 2460  
2401 TACAAATGGCAATCAATCTGAGATGCGGAGGAGCCAAAAAGTTTCAAGAAAGGCGA 2460  
2461 AAGGACATGACTCCATGTATGCAAGTCAATGAGAGACCAATGATATAGGAGATCTG 2520  
2461 AAGGACATGACTCCATGTATGCAAGTCAATGAGAGACCAATGATATAGGAGATCTG 2520  
2521 CTACAGAGATTCAGAGGAGCTCTTCTGTGAGCAGAGAGTGAACCTTACCGGCGTTCCAG 2580  
2521 CTACAGAGATTCAGAGGAGCTCTTCTGTGAGCAGAGAGTGAACCTTACCGGCGTTCCAG 2580  
2581 GGCACATAGGGGCTGTCTCTCTCCCAACCAATATGCTCCAGGGCCCAACTGCA 2640  
2581 GGCACATAGGGGCTGTCTCTCTCCCAACCAATATGCTCCAGGGCCCAACTGCA 2640  
2641 AAGTTGGGCACTGAGAGGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2700  
2641 AAGTTGGGCACTGAGAGGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2700  
2701 ACCTTCTCCATCCCAACAAATGGGATGTAAAGACAGACAGACATTCCTTACTG 2760

2701 ACCTTCTCCATCCCAACAAATGGGATGTAAAGACAGACAGACATTCCTTACTG 2760  
2761 AACACTGAGAGCCCATGAGACCCAGAGAAATTAATTGATTCATTCCAGACCTTGTGTA 2820  
2761 AACACTGAGAGCCCATGAGACCCAGAGAAATTAATTGATTCATTCCAGACCTTGTGTA 2820  
2821 GTTTCATTAAGCAGGACCTGAGACACCCGTCGGTTCCTTACCGAAATCTTAAGAA 2880  
2821 GTTTCATTAAGCAGGACCTGAGACACCCGTCGGTTCCTTACCGAAATCTTAAGAA 2880  
2881 GTTTCATTAAGCAGGACCTGAGACACCCGTCGGTTCCTTACCGAAATCTTAAGAA 2880  
2881 GTTTCATTAAGCAGGACCTGAGACACCCGTCGGTTCCTTACCGAAATCTTAAGAA 2880  
2941 TCAAGGCTCATTTCAAGAGGCAAGATTAAGAAATGATGAATTCATTCGATATACAGT 3000  
2941 TCAAGGCTCATTTCAAGAGGCAAGATTAAGAAATGATGAATTCATTCGATATACAGT 3000  
3001 CATGACAGCTCATGTGCTCTCAACTTAAAGCTGTGAGGAGCCGCTGATATAGAGAG 3060  
3001 CATGACAGCTCATGTGCTCTCAACTTAAAGCTGTGAGGAGCCGCTGATATAGAGAG 3060  
3061 AGAGAGCCTGAGTCACTTACATAGGCTGTGAGAGACCCCTGATTAAGTGTAAAC 3120  
3061 AGAGAGCCTGAGTCACTTACATAGGCTGTGAGAGACCCCTGATTAAGTGTAAAC 3120  
3121 AGAGGCTTGGCCCTGTGAGGCAAGATTAAGAAATGATGAATTCATTCGATATACAGT 3180  
3121 AGAGGCTTGGCCCTGTGAGGCAAGATTAAGAAATGATGAATTCATTCGATATACAGT 3180  
3181 CTCTCAGTGGGAGTCCCGAGATGAAGAAAGCAATGTGCTTTTATTTATTTATTTG 3240  
3181 CTCTCAGTGGGAGTCCCGAGATGAAGAAAGCAATGTGCTTTTATTTATTTATTTG 3240  
3241 TGCTCTGTGTATTTAAAGATCAATATTAACCACTTACTTCTTCACTGACTTA 3300  
3241 TGCTCTGTGTATTTAAAGATCAATATTAACCACTTACTTCTTCACTGACTTA 3300  
3301 GTATTAATCACTTACTTACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3360  
3301 GTATTAATCACTTACTTACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3360  
3361 AAACGTGCTGTGCTCCCGAGAGTGAAGAAATTTTCAATCTGTCCAAACCAAGAAAGA 3420  
3361 AAACGTGCTGTGCTCCCGAGAGTGAAGAAATTTTCAATCTGTCCAAACCAAGAAAGA 3420  
3421 ATGTGTGTGTGAGACAGATTTGACATATCTGTGATTAAGAGATTCCTGATTC 3480  
3421 ATGTGTGTGTGAGACAGATTTGACATATCTGTGATTAAGAGATTCCTGATTC 3480  
3481 TAGGTGCTGTGTGTGATTTCCATTTGTGAATTCATCTTGAATTCCTGATTC 3540  
3481 TAGGTGCTGTGTGTGATTTCCATTTGTGAATTCATCTTGAATTCCTGATTC 3540  
3541 CCTAGCAATTAAGAAATTTCTCAAGTTTCAATGTGAGGCTTCTCTGAGTCAATTA 3600  
3541 CCTAGCAATTAAGAAATTTCTCAAGTTTCAATGTGAGGCTTCTCTGAGTCAATTA 3600  
3601 CTGTGACATTTAAAGAGAAATTTAGAGAAATTTCTCATCTCTTAAATATTA 3660  
3601 CTGTGACATTTAAAGAGAAATTTAGAGAAATTTCTCATCTCTTAAATATTA 3660  
3661 TACCAAAAGAGGAGGCTCTGATTAATTTCTGTGAGCAGAGCAGCAGCAGCAGCAGCAG 3720  
3661 TACCAAAAGAGGAGGCTCTGATTAATTTCTGTGAGCAGAGCAGCAGCAGCAGCAGCAG 3720  
3721 CTTAATAAACAACATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 3780  
3721 CTTAATAAACAACATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 3780  
3781 GAATGAATAATCAAGGTGTCAAGAGCTGTGCTCTCTCTGAGAGCTTAAGAGAGAGCG 3840

Dh 3781 GAATGAAATCAAGGTCTAGACAGAGCTGTGCTCTTTCTTGAAAGGCTTGAAGGAAGCCG 3840  
Qy 3841 GTTCCTTGGCATTTCAAGCTTCTAGAGGCTGAGGCTGAGATTTCCAGAGCTCCAGTGGCTGTC 3900  
Dh 3841 GTTCCTTGGCATTTCAAGGCTTCTAGAGGCTGAGGCTGAGATTTCCAGAGCTCCAGTGGCTGTC 3900  
Qy 3901 AAGCTTTTCTCAGATGGCATCACTGTGACACTGAGGCTCCACTTCCCTTTGACTTAC 3960  
Dh 3901 AAGCTTTTCTCAGATGGCATCACTGTGACACTGAGGCTCCACTTCCCTTTGACTTAC 3960  
Qy 3961 AAAGCCACAGGAAGATCCAGGATATCTCTCATCTAAAGATCTTATCATCTCTGA 4020  
Dh 3961 AAAGCCACAGGAAGATCCAGGATATCTCTCATCTAAAGATCTTATCATCTCTGA 4020  
Qy 4021 AGAGCTTTTGGCATCAAGACAATAGCCAGGTGGGATTTAGAGCAGAGACATCTT 4080  
Dh 4021 AGAGCTTTTGGCATCAAGACAATAGCCAGGTGGGATTTAGAGCAGAGACATCTT 4080  
Qy 4081 TGGGCTGCTGTTATTTGCTTACACACCTTCTGCACTGACTCCCAAGAGAGCTA 4140  
Dh 4081 TGGGCTGCTGTTATTTGCTTACACACCTTCTGCACTGACTCCCAAGAGAGCTA 4140  
Qy 4141 CAAAATGATCTGGCGACAGGAGTGTGTTTGTAGCTGGAGCTTAAACATTAAAAA 4200  
Dh 4141 CAAAATGATCTGGCGACAGGAGTGTGTTTGTAGCTGGAGCTTAAACATTAAAAA 4200  
Qy 4201 ACCCCAGATCAGAAAGATCTGGCCATCTGGGCTCAATTTCTACCTTAGCAACATGCG 4260  
Dh 4201 ACCCCAGATCAGAAAGATCTGGCCATCTGGGCTCAATTTCTACCTTAGCAACATGCG 4260  
Qy 4261 TGGAGCTGGGACACAGCTCTGCTTTTGAAGAGGGGTGCCATTACCAAGTCAACAC 4320  
Dh 4261 TGGAGCTGGGACACAGCTCTGCTTTTGAAGAGGGGTGCCATTACCAAGTCAACAC 4320  
Qy 4321 CCACATAGGCCCTTACCTTCCACACAAGAGGCTAAGTGTGTTTCTTCTAGATCAATG 4380  
Dh 4321 CCACATAGGCCCTTACCTTCCACACAAGAGGCTAAGTGTGTTTCTTCTAGATCAATG 4380  
Qy 4381 CCCCTCAGGTTGATTTATTTGTAATGAAGAAAGAAAGAGCTGGATTAATCTTAATCAG 4440  
Dh 4381 CCCCTCAGGTTGATTTATTTGTAATGAAGAAAGAAAGAGCTGGATTAATCTTAATCAG 4440  
Qy 4441 TGAATGACCATGAGACCAATGTGTGCTACATTACCCTTTTCTTTTCTTTTCT 4500  
Dh 4441 TGAATGACCATGAGACCAATGTGTGCTACATTACCCTTTTCTTTTCTTTTCT 4500  
Qy 4501 TTTTCTTTTCTTTTAAATGTGAGAGAGATCTCATTTGCTGGCTAAGCTGAGTGCAG 4560  
Dh 4501 TTTTCTTTTCTTTTAAATGTGAGAGAGATCTCATTTGCTGGCTAAGCTGAGTGCAG 4560  
Qy 4561 TGGCGCAATCTCGGCTCACTGCAACCTCTGCTCTGGGCTCAAGCAATTTCTCCACTC 4620  
Dh 4561 TGGCGCAATCTCGGCTCACTGCAACCTCTGCTCTGGGCTCAAGCAATTTCTCCACTC 4620  
Qy 4621 AGCTCTCCAAATAGCTGGGATCACTGGCAAAACCAATGCCCACATAATTTTGTATTT 4680  
Dh 4621 AGCTCTCCAAATAGCTGGGATCACTGGCAAAACCAATGCCCACATAATTTTGTATTT 4680  
Qy 4681 TTTGTAGACAGAGGTTTCAACCAATGTGCGCAGGCTGTCTCAACCTCCCTGGGCTCAAC 4740  
Dh 4681 TTTGTAGACAGAGGTTTCAACCAATGTGCGCAGGCTGTCTCAACCTCCCTGGGCTCAAC 4740  
Qy 4741 AATCTCTGCTCGGCTCCCAAAAGTGTGAGATTACAGATGTGAGCCACGCACTCAG 4800  
Dh 4741 AATCTCTGCTCGGCTCCCAAAAGTGTGAGATTACAGATGTGAGCCACGCACTCAG 4800  
Qy 4801 CCCCAACCTCTATTAACCAATTAACCTGCGCAATGCTGAGACTTTTGTCTCTCAC 4860  
Dh 4801 CCCCAACCTCTATTAACCAATTAACCTGCGCAATGCTGAGACTTTTGTCTCTCAC 4860  
Qy 4861 CCTGTCTGATCTGGAAGAGAGGATTAATGTTAATAGCTTGTCAACACAGTCCCAAGTT 4920  
Dh 4861 CCTGTCTGATCTGGAAGAGAGGATTAATGTTAATAGCTTGTCAACACAGTCCCAAGTT 4920

Qy 4921 CAATATTTCTGGGCAAAAACCTTCTTCAAAAATTAATGTAATTTGATTTCAATGA 4980  
Dh 4921 CAATATTTCTGGGCAAAAACCTTCTTCAAAAATTAATGTAATTTGATTTCAATGA 4980  
Qy 4981 AATGACCTTGAAGAACCCGCTCACTGTGCAATGGCAATTAATGAAGAAATTTTA 5040  
Dh 4981 AATGACCTTGAAGAACCCGCTCACTGTGCAATGGCAATTAATGAAGAAATTTTA 5040  
Qy 5041 TAGTCTCTTAATGAGCGTGTACTGCAAGACCTTTGAACACTTTCCAGAGATAGATAT 5100  
Dh 5041 TAGTCTCTTAATGAGCGTGTACTGCAAGACCTTTGAACACTTTCCAGAGATAGATAT 5100  
Qy 5101 TTAAGTCAATGCTTGGCGTGTGCTATGAGCACTTTCCCTTGTAAAGTCTGCTCTGC 5160  
Dh 5101 TTAAGTCAATGCTTGGCGTGTGCTATGAGCACTTTCCCTTGTAAAGTCTGCTCTGC 5160  
Qy 5161 CCAAGTACCTTGGGCTTGTAGCCGGAATGTGACCTTGCAATAAGGGCCAAAGAGGG 5220  
Dh 5161 CCAAGTACCTTGGGCTTGTAGCCGGAATGTGACCTTGCAATAAGGGCCAAAGAGGG 5220  
Qy 5221 CTGGGCTTCTTCCCTCAGTGAAGAGCCCTTAATTTGAATTCATGCTGTGAGCCCTAGC 5280  
Dh 5221 CTGGGCTTCTTCCCTCAGTGAAGAGCCCTTAATTTGAATTCATGCTGTGTAGCCCTAGC 5280  
Qy 5281 CCTCAATCTGGAATTTCCCAACCTCCAGCCCTTCCAGACAGACTAGGTGCCCTGC 5340  
Dh 5281 CCTCAATCTGGAATTTCCCAACCTCCAGCCCTTCCAGACAGACTAGGTGCCCTGC 5340  
Qy 5341 ATTCAACCCAGGTGGGATTTGGCTTCTTGGGCTGCTGCTACTGTGACCATCAACGACAT 5400  
Dh 5341 ATTCAACCCAGGTGGGATTTGGCTTCTTGGGCTGCTGCTACTGTGACCATCAACGACAT 5400  
Qy 5401 CACTGTGCTGTGCAAGACACCAAGTGGCCATTTTCTTCAACTGAGGGCTCAAAACTCC 5460  
Dh 5401 CACTGTGCTGTGCAAGACACCAAGTGGCCATTTTCTTCAACTGAGGGCTCAAAACTCC 5460  
Qy 5461 TGAACAAGTGTGCTCTCTGAGACAGATTTCTGTAGCTGTGCTCTGAGTGAAGGGGC 5520  
Dh 5461 TGAACAAGTGTGCTCTCTGAGACAGATTTCTGTAGCTGTGCTCTGAGTGAAGGGGC 5520  
Qy 5521 CCAAGCTGAGGAACCTGCTCTTTTCTTAAAGCCAGGCCCACTTACATTAATAACAT 5580  
Dh 5521 CCAAGCTGAGGAACCTGCTCTTTTCTTAAAGCCAGGCCCACTTACATTAATAACAT 5580  
Qy 5581 TCAGGATCACTGGAACAAGTGAAGTGTGTAAGCTTCTGACATGACCCCACT 5640  
Dh 5581 TCAGGATCACTGGAACAAGTGAAGTGTGTAAGCTTCTGACATGACCCCACT 5640  
Qy 5641 GCTCATCAAGTGTGCTGCTGATGCTTACGAGAGAGCCAGCGCATGACAGACTGTCTCT 5700  
Dh 5641 GCTCATCAAGTGTGCTGCTGATGCTTACGAGAGAGCCAGCGCATGACAGACTGTCTCT 5700  
Qy 5701 AATGCTGTGTATTTGACACAGAGGAAAGGCTCAAGAGAAAGTCACTGGAGCAAGCA 5760  
Dh 5701 AATGCTGTGTATTTGACACAGAGGAAAGGCTCAAGAGAAAGTCACTGGAGCAAGCA 5760  
Qy 5761 CAAGCCACCCGACATGAGCTTGTGAAGGTTAGAGAGCTGTGTGTGATCTGCACT 5820  
Dh 5761 CAAGCCACCCGACATGAGCTTGTGAAGGTTAGAGAGCTGTGTGTGATCTGCACT 5820  
Qy 5821 GCTTCACTGGAATTAATTAATTTCAATGAGATCTTTTAAAGTGCATTTTAATTCATTT 5880  
Dh 5821 GCTTCACTGGAATTAATTAATTTCAATGAGATCTTTTAAAGTGCATTTTAATTCATTT 5880  
Qy 5881 CTGTGCTTTAATAAACAAGTATACCAAAAAACAAGTATCAAGCTGTTTAAGTGTGG 5940  
Dh 5881 CTGTGCTTTAATAAACAAGTATACCAAAAAACAAGTATCAAGCTGTTTAAGTGTGG 5940  
Qy 5941 CTACTGTGCTGCTGTGATGAGTGAAGCCCGGTTTCCAGTGTGTAAGTGAAGGCTC 6000  
Dh 5941 CTACTGTGCTGCTGTGATGAGTGAAGCCCGGTTTCCAGTGTGTAAGTGAAGGCTC 6000

QY 6001 ACATGGGCTCAGACAGATCTGTCTTAATTGTGATGATACAGAAAGCCAGGCTTTGGG 6060  
DB 6001 ACATGGGCTCAGACAGATCTGTCTTAATTGTGATGATACAGAAAGCCAGGCTTTGGG 6060  
QY 6061 ATACAGATCTTCTCTCTTCAATTTGATGCGCGTCACTGTGTGAAGAGATGTTTTTGTCC 6120  
DB 6061 ATACAGATCTTCTCTCTTCAATTTGATGCGCGTCACTGTGTGAAGAGATGTTTTTGTCC 6120  
QY 6121 GGAATATATATATATAGTCTTGAGTCTGCGCCAAAAA 6163  
DB 6121 GGAATATATATATATAGTCTTGAGTCTGCGCCAAAAA 6163

RESULT 2  
AY026461 5963 bp mRNA linear PRI 17-AUG-2001  
LOCUS Homo sapiens CUB domain containing protein 1 (CDCP1) mRNA, complete  
DEFINITION  
ACCESSION AY026461  
VERSION AY026461.1 GI:14328878  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 5963)  
Scherl-Mostagguer, W., Sommergruber, W., Abseher, R., Hauptmann, R.,  
Ambros, P. and Schweifer, N.  
Identification of a novel gene, CDCP1, overexpressed in human  
colorectal cancer  
Oncogene 20 (32), 4402-4408 (2001)  
PUBMED 11466621  
TITLE 2 (bases 1 to 5963)  
Scherl-Mostagguer, W., Sommergruber, W., Abseher, R., Hauptmann, R.,  
Ambros, P. F. and Schweifer, N.  
Direct Submision  
JOURNAL Submitted (31-JAN-2001) Exploratory Research, Boehringer Ingelheim,  
Austria, Dr. Boehringergasse 5-11, Vienna 1121, Austria

FEATURES  
source  
1..5963  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="3"  
/map="3p23-p21"  
/cell\_line="A549; Calu 6"  
1..5963  
/gene="CDCP1"  
1..82  
/gene="CDCP1"  
83..2593  
/gene="CDCP1"  
/note="transmembrane protein"  
/codon\_start=1  
/product="CUB domain containing protein 1"  
/protein\_id="AAK02058.1"  
/db\_xref="GI:14328878"  
/translation="MAGLNGVSIALLGVLLGAARLPRGAAFEIALPRESNITVLI  
KLGPTPLAKPYIVISKRHITMLSIKSGERIVFTFSCSPENHVIIELOKXIDCMG  
PPGPEVQLQPSLPTLNRFTIMDKVSKSIGLELOISPRLOIGSGESCGSVT  
HSISGRIDATVIRIGFTGNGTVSRIKMOEGKMLHLHMPHRYNVSISNRISIK  
RUCITESVREGSATIMSNANPEGPPELMTQFVYAHARASVSPINPLNSCER  
KEERBYIIPGSTTNPPEVKLKDQGNMAGNPNLSLQCDDDAGSPGILRLQFVLY  
QHPQNSNKIIVYVLSNERAMSLTEPRVVKOSRKFVPCFVCLSRSTSSLTLSG  
SHKISFLCDDLTRLMNVYEKTSCTDHRVQCKSISLQSPDIIHLPELHDFSMKG  
LVPKRLSLVLPAAQLQHTHEKPCNTSFYSIVASAIQSPDLRGSFCPGSIIQIQ  
VYONISVTLRTFAPSPFOGASROGLVSPFIPYKEGAVPTVMDTKSKYLRTPMMDR  
GIPSLTSVSNMISVPRDOVACLTPFERSGAVICOTGRAMIIQDQRTAEELPSLDED  
VLPKPSFHHHSFVNI SNCSPTSGKOLDLFSYTLTPRVDLTIVLIIAAVGGVLLS  
ALGLIICVKKKKKTKNKPKPAVGITNGINTEMPTPKPFQGRKNDKSHVAIVEDT  
NVYGHLDSSGSLFLOPEVDIYRPFQGTGVCPPSPPTCSAPAKTAKATEBP PPSPP  
PESEBEPYTFSHPNNGVSSKTDIDIPLNTQEMEPAE"

3'UTR 2594..5963  
BASE COUNT 1463 a 1634 c 1392 g 1474 t  
ORIGIN  
Query Match 96.8%; Score 5963; DB 9; Length 5963;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 5963; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 201 GCGACCTGGCGCGGCTTGGGGGCGCTGGGGGCGCGCTCCCGACGCTGTTCCCAAC 260  
DB 1 GCGACCTGGCGCGGCTTGGGGGCGCTGGGGGCGCGCTCCCGACGCTGTTCCCAAC 60  
QY 261 GAGGCGGAGGCGTCCCGGAGTCATGACCGGCTGAACTGCGGGGTCTGTATGCACTGCT 320  
DB 61 GAGGCGGAGGCGTCCCGGAGTCATGACCGGCTGAACTGCGGGGTCTGTATGCACTGCT 120  
QY 321 AGGGGTTCTGCTGCTGGGGTGGCGCGCTGCGCGCGCGGAGCAAGCTTTGAATTC 380  
DB 121 AGGGGTTCTGCTGCTGGGGTGGCGCGCTGCGCGCGCGGAGCAAGCTTTGAATTC 180  
QY 381 TCTGCGACGAGAAAGCAATTACAGTTCTCAATAAGCTGGGAGACCCGACCTCTGCTGGC 440  
DB 181 TCTGCGACGAGAAAGCAATTACAGTTCTCAATAAGCTGGGAGACCCGACCTCTGCTGGC 240  
QY 441 AAAACCTGTTACATGTCATTTCTAAAGACATATAACATGTTGTCATCAAGTCTGG 500  
DB 241 AAAACCTGTTACATGTCATTTCTAAAGACATATAACATGTTGTCATCAAGTCTGG 300  
QY 501 AGAAAGAAATAGTCTTTACCTTTAGCTGCGCAGAGTCTTAAGAAATCACTTTGTATAGAT 560  
DB 301 AGAAAGAAATAGTCTTTACCTTTAGCTGCGCAGAGTCTTAAGAAATCACTTTGTATAGAT 360  
QY 561 CCAGAAATATATGACTGATGATGACGCCCAAGTCTTTGGGGAGTTCACCTTCAGGC 620  
DB 361 CCAGAAATATATGACTGATGATGACGCCCAAGTCTTTGGGGAGTTCACCTTCAGGC 420  
QY 621 CTGACATGCTGTTGGCTCAACCTCAACAGAACTTTCACTTGGAGTCAAAAGTCAATTA 680  
DB 421 CTGACATGCTGTTGGCTCAACCTCAACAGAACTTTCACTTGGAGTCAAAAGTCAATTA 480  
QY 681 GAGCATCGGTTTGAAGCTGCACTGATTTTCATCCCTGCTGAGGAGATCGTGGGTTGA 740  
DB 481 GAGCATCGGTTTGAAGCTGCACTGATTTTCATCCCTGCTGAGGAGATCGTGGGTTGA 540  
QY 741 GAGCTGCCCAAGCGAGTCACTCATCTCATGACGCGCGGAATGCAATGCCACCTGTGAG 800  
DB 541 GAGCTGCCCAAGCGAGTCACTCATCTCATGACGCGCGGAATGCAATGCCACCTGTGAG 600  
QY 801 GATCGGAACCTTCTGACAGATGACACTGTGTCCCGGATCAAGATGCAAGAGAGTGA 860  
DB 601 GATCGGAACCTTCTGACAGATGACACTGTGTGTCCCGGATCAAGATGCAAGAGAGTGA 660  
QY 861 AATGGCTTTACACTTCAGAGTTCACCCGAGAAATGCTTCGCGCTTCAGATTGCAA 920  
DB 661 AATGGCTTTACACTTCAGAGTTCACCCGAGAAATGCTTCGCGCTTCAGATTGCAA 720  
QY 921 CCGCTCATCTATAAAGCTGTGTGATCATGCAAGTGTGTTGAGGGTGAAGGCTCAGC 980  
DB 721 CCGCTCATCTATAAAGCTGTGTGATCATGCAAGTGTGTTGAGGGTGAAGGCTCAGC 780  
QY 981 AACCTGATGTGCACTACCAAGAAAGGCTTCCTGAGATGAGTCAATGAGTGA 1040  
DB 781 AACCTGATGTGCACTACCAAGAAAGGCTTCCTGAGATGAGTCAATGAGTGA 840  
QY 1041 GTTGTGCTTCTGTGCAACTGCGGAGCGCTCTCTCTCAACTTCACCTCTGCAA 1100  
DB 841 GTTGTGCTTCTGTGCAACTGCGGAGCGCTCTCTCTCAACTTCACCTCTGCAA 900  
QY 1101 CTGTGAGAGAAAGAGAGCGGGTGAATCTAATCTCCGGGCTCCACCAACCCGGA 1160  
DB 901 CTGTGAGAGAAAGAGAGCGGGTGAATCTAATCTCCGGGCTCCACCAACCCGGA 960

QY 1161 GGTGTTCAAGCTGGAGACAGACGCTGGGAAATGCGGGGAACTTCAACTCTCTCT 1220  
Db 961 GGTTTCAAGCTGGAGACAGACGCTGGGAAATGCGGGGAACTTCAACTCTCTCT 1020  
QY 1221 GCAAGGCTGGACCAAGATGCCAAGTCCAGGGATCTTCGGCTGAGTCCAGTTT 1280  
Db 1021 GCAAGGCTGGACCAAGATGCCAAGTCCAGGGATCTTCGGCTGAGTCCAGTTT 1080  
QY 1281 GGTCCAAATCAAAATGAAAGCAATAAATCTAGTGTGACTTGAATGATGCG 1340  
Db 1081 GGTCCAAATCAAAATGAAAGCAATAAATCTAGTGTGACTTGAATGATGCG 1140  
QY 1341 AGCAGTCTCACTCAACCATGAGCCAGGCGCCCTCAAAAGAGCCGCAAGTTTCCCTG 1400  
Db 1141 AGCAGTCTCACTCAACCATGAGCCAGGCGCCCTCAAAAGAGCCGCAAGTTTCCCTG 1200  
QY 1401 CTGTTTCTGATCTGGAATCTGGACCTGAGTGAACCTGACCTGATCATCTGGCTC 1460  
Db 1201 CTGTTTCTGATCTGGAATCTGGACCTGAGTGAACCTGACCTGATCATCTGGCTC 1260  
QY 1461 CAACAACAAATCTCTTCTTTGTGATGATCTGACACGTCTGTGATGATGAAAA 1520  
Db 1261 CAACAACAAATCTCTTCTTTGTGATGATCTGACACGTCTGTGATGATGAAAA 1320  
QY 1521 AACCATTAAGTGCACAGACCAAGGATCTGACCAAGGAAATCTTACTCACTCAAGTGC 1580  
Db 1321 AACCATTAAGTGCACAGACCAAGGATCTGACCAAGGAAATCTTACTCACTCAAGTGC 1380  
QY 1581 CAGTGAATCTCACTGCTGCTGAGTGTGAGTGTGACATGACTTCTCTGGAGCTGTGGTGC 1640  
Db 1381 CAGTGAATCTCACTGCTGCTGAGTGTGAGTGTGACATGACTTCTCTGGAGCTGTGGTGC 1440  
QY 1641 CAAGGACAGGCTCAGCTGTGTGTGTGTCAGGCCAGAAAGCTGACAGACATACACGA 1700  
Db 1441 CAAGGACAGGCTCAGCTGTGTGTGTGTCAGGCCAGAAAGCTGACAGACATACACGA 1500  
QY 1701 GAAGCCTGCAACACAGCTTCACTGCTGTGACAGTGCATACCAAGCCAGGACCT 1760  
Db 1501 GAAGCCTGCAACACAGCTTCACTGCTGTGACAGTGCATACCAAGCCAGGACCT 1560  
QY 1761 GTACTTCTGCTCTTCTGCGCGGGAGGCTCTATCAAGCAGATTCAGGTGAAGCAGAACT 1820  
Db 1561 GTACTTCTGCTCTTCTGCGCGGGAGGCTCTATCAAGCAGATTCAGGTGAAGCAGAACT 1620  
QY 1821 CTGCTGACCTTTCGACCTTTTGGCCCACTTTCACAAAGAGGCTCCAGGAGGCTCT 1880  
Db 1621 CTGCTGACCTTTCGACCTTTTGGCCCACTTTCACAAAGAGGCTCCAGGAGGCTCT 1680  
QY 1881 GACGGTGTCTTTATACCTTATTTCAAAAGAGAGGCTTTTCAAGGTGACCTTGAC 1940  
Db 1681 GACGGTGTCTTTATACCTTATTTCAAAAGAGAGGCTTTTCAAGGTGACCTTGAC 1740  
QY 1941 AAAAAGCAAGTCTACCTGAGGACCCCAACTGGGACCGGGGCTGCACTCCCTCACTC 2000  
Db 1741 AAAAAGCAAGTCTACCTGAGGACCCCAACTGGGACCGGGGCTGCACTCCCTCACTC 1800  
QY 2001 TGTGCTCTGAGACATCAGCTGTGACAGAGACAGGTGCTGCTGCTTCTTTAAGA 2060  
Db 1801 TGTGCTCTGAGACATCAGCTGTGACAGAGACAGGTGCTGCTGCTTCTTTAAGA 1860  
QY 2061 GCGGAGCGGCTGTGCTGACAGAGAGGCGGCAATTCATGATCATCCAGAGAGCGGAC 2120  
Db 1861 GCGGAGCGGCTGTGCTGACAGAGAGGCGGCAATTCATGATCATCCAGAGAGCGGAC 1920  
QY 2121 CCGGGCTGAGAGATCTTCAAGCTGAGAGAGAGTGTGCTCCCAAGGCAAGTCTTCAACA 2180  
Db 1921 CCGGGCTGAGAGATCTTCAAGCTGAGAGAGAGTGTGCTCCCAAGGCAAGTCTTCAACA 1980  
QY 2181 TCACAGCTTCTGGGTCAACATCTTAACTGACAGCCCAAGAGCGGCAAGAGCTAAGCT 2240  
Db 1981 TCACAGCTTCTGGGTCAACATCTTAACTGACAGCCCAAGAGCGGCAAGAGCTAAGCT 2040  
QY 2241 GCTCTTCTCGGTGACCTTACCCCAAGAGCTGTGACTTGAATCTCTCATCGGACG 2300

Db 2041 GCTCTTCTCGGTGACCTTACCCCAAGAGCTGTGACTTGAATCTCTCATCGGACG 2100  
QY 2301 GGTGGAGGTGAGTCTTACTGCTGTGACCTCGGGCTCATCATTTGCTGTGAAAA 2360  
Db 2101 GGTGGAGGTGAGTCTTACTGCTGTGACCTCGGGCTCATCATTTGCTGTGAAAA 2160  
QY 2361 GAAAGAAAAAGAACAAAGAGGCGCGCTGTGGATCTTAATGGAACATCAATAC 2420  
Db 2161 GAAAGAAAAAGAACAAAGAGGCGCGCTGTGGATCTTAATGGAACATCAATAC 2220  
QY 2421 TGAGATCCGAGGCGCAAAAGTTTCAGAAAGGCGGAAAGGCAATGACTCCCATGT 2480  
Db 2221 TGAGATCCGAGGCGCAAAAGTTTCAGAAAGGCGGAAAGGCAATGACTCCCATGT 2280  
QY 2481 GTATGAGTCACTGAGAGACACCATGTATATGGGATCTGTACAGATTTCCAGCGCTC 2540  
Db 2281 GTATGAGTCACTGAGAGACACCATGTATATGGGATCTGTACAGATTTCCAGCGCTC 2340  
QY 2541 CTTCCTGACCGAGAGGTGACACCTTACCGGCTTCCAGGCGACCATGGGGTCTGTCC 2600  
Db 2341 CTTCCTGACCGAGAGGTGACACCTTACCGGCTTCCAGGCGACCATGGGGTCTGTCC 2400  
QY 2601 TCCCTCCCAACCACTATGTGTCCAGGGCGCCCACTGAAAGTTGGCCCATGAGAGCC 2660  
Db 2401 TCCCTCCCAACCACTATGTGTGTCCAGGGCGCCCACTGAAAGTTGGCCCATGAGAGCC 2460  
QY 2661 ACCCTCTCGCTCCCTCTCTGAGTGTGAGTGAACCGTACACTTCTCCATCCCAACA 2720  
Db 2461 ACCCTCTCGCTCCCTCTCTGAGTGTGAGTGAACCGTACACTTCTCCATCCCAACA 2520  
QY 2721 TGGGATGTAAAGCAGAGACACAGACATTCCTTACTGAAACACTCAGAGGCCATGA 2780  
Db 2521 TGGGATGTAAAGCAGAGACACAGACATTCCTTACTGAAACACTCAGAGGCCATGA 2580  
QY 2781 GCAGAGAAATTAATCTGATTCATCCAGAGGCTTGTGAGTTTCAATAAGAGGCACT 2840  
Db 2581 GCAGAGAAATTAATCTGATTCATCCAGAGGCTTGTGAGTTTCAATAAGAGGCACT 2640  
QY 2841 GAGACACCGCTCGGTCTCTTAACCAAGAAATCTTAAGAGAGAAATTAACAGAGAA 2900  
Db 2641 GAGACACCGCTCGGTCTCTTAACCAAGAAATCTTAAGAGAGAAATTAACAGAGAA 2700  
QY 2901 CAGCAGAGGTTTTCTTGACACCGCAACTTCACTGCTGATGACTCATTTAAG 2960  
Db 2701 CAGCAGAGGTTTTCTTGACACCGCAACTTCACTGCTGATGACTCATTTAAG 2760  
QY 2961 GCAAGCATTTGAATGATGAATTCGAATCTGATGATGATGATGATGATGATGATG 3020  
Db 2761 GCAAGCATTTGAATGATGAATTCGAATCTGATGATGATGATGATGATGATGATG 2820  
QY 3021 TCACATTTAGCTGTGCGTTTACAGCTGTATGAGAGAGAGGCTGATCACTTA 3080  
Db 2821 TCACATTTAGCTGTGCGTTTACAGCTGTATGAGAGAGAGGCTGATCACTTA 2880  
QY 3081 GCATAGGTTTGAAGAGCGCTTGATTCAGAGTGTAAACAGAGGCTTCCCTTCAAG 3140  
Db 2881 GCATAGGTTTGAAGAGCGCTTGATTCAGAGTGTAAACAGAGGCTTCCCTTCAAG 2940  
QY 3141 ACAACGTTTCAATTCAGAGGCTTACCTGAGTCCCTCACTCTGAGGCTCCCAAG 3200  
Db 2941 ACAACGTTTCAATTCAGAGGCTTACCTGAGTCCCTCACTCTGAGGCTCCCAAG 3000  
QY 3201 ATGAAAACGACATGCTCTTTTATTTATTTATTTATTTATTTATTTATTTATTTAAG 3260  
Db 3001 ATGAAAACGACATGCTCTTTTATTTATTTATTTATTTATTTATTTATTTATTTAAG 3060  
QY 3261 GATCAATATATTAACCACTGATCTTTTCACTGACCTTATATTAATCTATTAATCT 3320  
Db 3061 GATCAATATATTAACCACTGATCTTTTCACTGACCTTATATTAATCTATTAATCT 3120  
QY 3321 GTTGGATGCTGTGGTTGATCTTCACTGACCGCTAATTAACGTTGCTGTCCCA 3380

Db 3121 GTTTGATGCTGGGTTGTGACTTCTAGACCGGTAGATAAAGTGTCCTGTCCCA 3180  
Qy 3381 GGTGTGGGAATTAATTAACAATCTGTCAACAGAAAGAAATGTGTGTTGAGCAGA 3440  
Db 3181 GGTGTGGGAATTAATTAACAATCTGTCAACAGAAAGAAATGTGTGTGAGCAGA 3240  
Qy 3441 TTGACACATATCTGCTTTGATTAAGAGACTTCTGATTTCTTAGTGGTTGTGTTATC 3500  
Db 3241 TTGACACATATCTGCTTTGATTAAGAGACTTCTGATTTCTTAGTGGTTGTGTTATC 3300  
Qy 3501 CCATTGTGAAAATTCATCTTGAATGCCATTTGCTTATGCTTCAAGAAATTAAGAAATTT 3560  
Db 3301 CCATTGTGAAAATTCATCTTGAATGCCATTTGCTTATGCTTCAAGAAATTAAGAAATTT 3360  
Qy 3561 CCTCAAGTTTCCATGTGCGGTTCTCCTAGCTGAGCAATACTTGAATTTAAAGAGAA 3620  
Db 3361 CCTCAAGTTTCCATGTGCGGTTCTCCTAGCTGAGCAATACTTGAATTTAAAGAGAA 3420  
Qy 3621 TTTAGAGAAATTTCTCATCTCTTAAATGTTTAAATATATACAAACAGTGGCCCTG 3680  
Db 3421 TTTAGAGAAATTTCTCATCTCTTAAATGTTTAAATATATACAAACAGTGGCCCTG 3480  
Qy 3681 CATTAAGTTTCTGTGTCCTGCACTGCAACCAATTAATCTTGTAGCTTTAAACAATTAAC 3740  
Db 3481 CATTAAGTTTCTGTGTCCTGCACTGCAACCAATTAATCTTGTAGCTTTAAACAATTAAC 3540  
Qy 3741 TTATAGTCTGTGGGATCAGAAATTCAAAAATGATGTCCTGAAATGAAATTCAGAGGTGA 3800  
Db 3541 TTATAGTCTGTGGGATCAGAAATTCAAAAATGATGTCCTGAAATGAAATTCAGAGGTGA 3600  
Qy 3801 GGAGAGCTGTGCTCTCTTGAAGGCTCTAGGAGAAAGCGGTTCTTGTCCATTTCAAGCT 3860  
Db 3601 GGAGAGCTGTGCTCTCTTGAAGGCTCTAGGAGAAAGCGGTTCTTGTCCATTTCAAGCT 3660  
Qy 3861 TCTAGAGGCTGTGCTGATTTCCAGAGGCTCCAGTGGCTGTGCAAGCTTTTCTCAGTGGCAT 3920  
Db 3661 TCTAGAGGCTGTGCTGATTTCCAGAGGCTCCAGTGGCTGTGCAAGCTTTTCTCAGTGGCAT 3720  
Qy 3921 CACTGTGACACTGGCCCTCCCACTTCTTGAATTAAGAAAGCCACAGGAAGATCC 3980  
Db 3721 CACTGTGACACTGGCCCTCCCACTTCTTGAATTAAGAAAGCCACAGGAAGATCC 3780  
Qy 3981 AGGATTAATCTCTCAATCTAAAGATCTTGTATCATCTCTGAAAGGCTTTTGTGCATGCAAG 4040  
Db 3781 AGGATTAATCTCTCAATCTAAAGATCTTGTATCATCTCTGAAAGGCTTTTGTGCATGCAAG 3840  
Qy 4041 ACAACCTAGCCACAGGTGGGATTAAGACACAGGACATCTTTGGGCTGTGTTATTTGCC 4100  
Db 3841 ACAACCTAGCCACAGGTGGGATTAAGACACAGGACATCTTTGGGCTGTGTTATTTGCC 3900  
Qy 4101 TACCAACCTTCTGCTGACATGACTCCACAGGAGAGGCTACAAATGATCTGGCCGACAG 4160  
Db 3901 TACCAACCTTCTGCTGACATGACTCCACAGGAGAGGCTACAAATGATCTGGCCGACAG 3960  
Qy 4161 GGATGTTTTTTAGCTTGGGACTTAACATTAATAAAACCCAGATCAGAAATCTG 4220  
Db 3961 GGATGTTTTTTAGCTTGGGACTTAACATTAATAAAACCCAGATCAGAAATCTG 4020  
Qy 4221 GGCATGTCTGGGCTCATTTCTACCTAGCAACATGTGCTGAGGCTGGCCACAGCTCT 4280  
Db 4021 GGCATGTCTGGGCTCATTTCTACCTAGCAACATGTGCTGAGGCTGGCCACAGCTCT 4080  
Qy 4281 GGCCTTTAGAAAGGGGTGCATCTTCAACAGGTGACCAAGCCCACTAGGCCCTTCACT 4340  
Db 4081 GGCCTTTAGAAAGGGGTGCATCTTCAACAGGTGACCAAGCCCACTAGGCCCTTCACT 4140  
Qy 4341 TCCCAAAATGAGGCTAAGTGTGTTTCTAATGATCAATGCCCCCTGACAGTGTCAATTAAT 4400  
Db 4141 TCCCAAAATGAGGCTAAGTGTGTTTCTAATGATCAATGCCCCCTGACAGTGTCAATTAAT 4200  
Qy 4401 TGTAAATGAAAAAGAAAGACTGGGATTAATCTCTAATCAGGTAGTAGACATGAGCCAA 4460  
Db 4201 TGTAAATGAAAAAGAAAGACTGGGATTAATCTCTAATCAGGTAGTAGACATGAGCCAA 4260

Qy 4461 TGTGTGCTCACAATACCCCTTTTCTTTTTTCTTTTTCTTTTTCTTTTTTAAAG 4520  
Db 4261 TGTGTGCTCACAATACCCCTTTTCTTTTTTCTTTTTCTTTTTTAAAG 4320  
Qy 4521 TGAACAGGATCTCAATTCCTGTTGCTTGAAGCTGAGTGAAGTGGCAATCTCGGCTACT 4580  
Db 4321 TGAACAGGATCTCAATTCCTGTTGCTTGAAGCTGAGTGAAGTGGCAATCTCGGCTACT 4380  
Qy 4581 GCAACCTGCTGCTCTGAGGCTCAAGCAATTCCTCCACCTGACCTCCAAATAGCTGGGA 4640  
Db 4381 GCAACCTGCTGCTCTGAGGCTCAAGCAATTCCTCCACCTGACCTCCAAATAGCTGGGA 4440  
Qy 4641 TCACTGGCAAAACCAACCAATGAGGCTCAATTTGATTTTGTAGAGACAGGTTTCA 4700  
Db 4441 TCACTGGCAAAACCAACCAATGAGGCTCAATTTGATTTTGTAGAGACAGGTTTCA 4500  
Qy 4701 CCATGTTGCCAGGCTGTGCTCAACCTCTGAGGCTCAAGCAATTCCTGCTCGGCTC 4760  
Db 4501 CCATGTTGCCAGGCTGTGCTCAACCTCTGAGGCTCAAGCAATTCCTGCTCGGCTC 4560  
Qy 4761 CCAAAGTGTGGGATTAAGATGAGGCAACCGCATCCAGCCCAACCTCATTTATAC 4820  
Db 4561 CCAAAGTGTGGGATTAAGATGAGGCAACCGCATCCAGCCCAACCTCATTTATAC 4620  
Qy 4821 CAATTAACCTGACCAATCTGGAATTTGCTTCTGCTCAACCTGCTGATCTGGAAG 4880  
Db 4621 CAATTAACCTGACCAATCTGGAATTTGCTTCTGCTCAACCTGCTGATCTGGAAG 4680  
Qy 4881 AGAGGATTAATGTTATGCTTGTGAGCAAGTCCCAAGTTCAATTTCTGGGCAAAA 4940  
Db 4681 AGAGGATTAATGTTATGCTTGTGAGCAAGTCCCAAGTTCAATTTCTGGGCAAAA 4740  
Qy 4941 CTTCTTCAAAAAATTAATGTTATCTTCAATTTGATTAAGATTTCACTTTGAAATGCAAC 5000  
Db 4741 CTTCTTCAAAAAATTAATGTTATCTTCAATTTGATTAAGATTTCACTTTGAAATGCAAC 4800  
Qy 5001 GCTTCAACCTGTTCAATGAGGCAATTAAGAAAGATTTATGCTTCTCAAAATGAGCGT 5060  
Db 4801 GCTTCAACCTGTTCAATGAGGCAATTAAGAAAGATTTATGCTTCTCAAAATGAGCGT 4860  
Qy 5061 ACTGCAAGACCTCTTGAACACTTTTCCAGAGATGAGATTAATTAATGATGCTTGGGCT 5120  
Db 4861 ACTGCAAGACCTCTTGAACACTTTTCCAGAGATGAGATTAATTAATGATGCTTGGGCT 4920  
Qy 5121 TGCCTATGACCTTTCCCTTCTGAAGTCTGCTTCTGAGGCTGACCTTGGCTGT 5180  
Db 4921 TGCCTATGACCTTTCCCTTCTGAAGTCTGCTTCTGAGGCTGACCTTGGCTGT 4980  
Qy 5181 GAGCCGAGATGCTGACCTTGAATTAAGGAGCAAGAGGAGGCTTCTTCCCTGAC 5240  
Db 4981 GAGCCGAGATGCTGACCTTGAATTAAGGAGCAAGAGGAGGCTTCTTCCCTGAC 5040  
Qy 5241 TGAAGAGCCCTTAATTTGAATTAATCTGTTGAGAGCCCTGACCTCAATCTGCAATTTCC 5300  
Db 5041 TGAAGAGCCCTTAATTTGAATTAATCTGTTGAGAGCCCTGACCTCAATCTGCAATTTCC 5100  
Qy 5301 CAACCTCCAGCCCTTCCAGAGAGGATGAGTGTCTGCAATTTCAACCAAGTGGGATTT 5360  
Db 5101 CAACCTCCAGCCCTTCCAGAGAGGATGAGTGTCTGCAATTTCAACCAAGTGGGATTT 5160  
Qy 5361 GGCCTTCTTGAAGGCTGCTGATCTGTCACATCAACGAGATCACTGTTGCTGCAAGGACA 5420  
Db 5161 GGCCTTCTTGAAGGCTGCTGATCTGTCACATCAACGAGATCACTGTTGCTGCAAGGACA 5220  
Qy 5421 CCAGTGGCAATTTTCTTCAACTGAGGCTCAAAATCTCTGCAATGTTGCTGCTCT 5480  
Db 5221 CCAGTGGCAATTTTCTTCAACTGAGGCTCAAAATCTCTGCAATGTTGCTGCTCT 5280  
Qy 5481 GAGACCAATTTTCTTGAAGCTGTGCTCAAGTGAAGGAGCCCAAGCTGAGGAACTCTGGC 5540  
Db 5281 GAGACCAATTTTCTTGAAGCTGTGCTCAAGTGAAGGAGCCCAAGCTGAGGAACTCTGGC 5340



QY 5541 TCTTTTCTTAAAGCCCGCCCACTTACATATAAACTTTCAAGGCTCACTGGAACAGT 5600  
DB 5341 TCTTTTCTTAAAGCCCGCCCACTTACATATAAACTTTCAAGGCTCACTGGAACAGT 5400  
QY 5601 GAAGTGCATTTTGTGAAGCTACTGATGCGACGCCCACTGCTCATCAAGTGTGCTGCC 5660  
DB 5401 GAAGTGCATTTTGTGAAGCTACTGATGCGACGCCCACTGCTCATCAAGTGTGCTGCC 5460  
QY 5661 ATGCTTACGAGGAAGCCGACGATGAGAGACTGCTCTTAATGCTGTGCTCATTTGCA 5720  
DB 5461 ATGCTTACGAGGAAGCCGACGATGAGAGACTGCTCTTAATGCTGTGCTCATTTGCA 5520  
QY 5721 GAAGGGAAGGCTCTCAAGGAAGAGTCACTGCGACCAAGCCCACTGCGACATAGGCC 5780  
DB 5521 GAAGGGAAGGCTCTCAAGGAAGAGTCACTGCGACCAAGCCCACTGCGACATAGGCC 5580  
QY 5781 TTGCTTAAAGGTTAGCAGACTGCTGTGTGATCTGCAAGTCTTCACTGGAATTAATTTA 5840  
DB 5581 TTGCTTAAAGGTTAGCAGACTGCTGTGTGATCTGCAAGTCTTCACTGGAATTAATTTA 5640  
QY 5841 TTGCTTAAAGGTTAGCAGACTGCTGTGTGATCTGCAAGTCTTCACTGGAATTAATTTA 5900  
DB 5641 TTGCTTAAAGGTTAGCAGACTGCTGTGTGATCTGCAAGTCTTCACTGGAATTAATTTA 5700  
QY 5901 TGTACCAAAAAACAAGTATCAAGCTGTTAAGCTTCCGCTACTTCCCTGCTTCAAG 5960  
DB 5701 TGTACCAAAAAACAAGTATCAAGCTGTTAAGCTTCCGCTACTTCCCTGCTTCAAG 5760  
QY 5961 TTAGAGGCGCGGTTTCCAGTGTGATGCTGACAGGCTCAGAGTGGGCTCAGCAGATGC 6020  
DB 5761 TTAGAGGCGCGGTTTCCAGTGTGATGCTGACAGGCTCAGAGTGGGCTCAGCAGATGC 5820  
QY 6021 TGTCTTAATTGTGATGATACAGAAAGCCAGGCTTGGGATACAAGTCTTCTCTTTC 6080  
DB 5821 TGTCTTAATTGTGATGATACAGAAAGCCAGGCTTGGGATACAAGTCTTCTCTTTC 5880  
QY 6081 ATTGTGATCCGCTGACCTGTGTGAAGCAGATGTTTTGTCCGGAATTAATTAATGTCT 6140  
DB 5881 ATTGTGATCCGCTGACCTGTGTGAAGCAGATGTTTTGTCCGGAATTAATTAATGTCT 5940  
QY 6141 TGGAGTCTGCGCAAAAAAAA 6163  
DB 5941 TGGAGTCTGCGCAAAAAAAA 5963

RESULT 3  
AX353603 5897 bp DNA linear PAT 06-FEB-2002  
LOCUS AX353603  
DEFINITION Sequence 1 from Patent. W00204508.  
ACCESSION AX353603.1 GI:18618676  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
AUTHORS 1  
TITLE Schwaefler, N., Scheerl-Mostagaer, M., Sommergruber, W., and Abseher, R.  
JOURNAL Tumour-associated antigen (b345), characterised by an amino acid  
FEATURES  
source  
location/Qualifiers  
1. 5897  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
1. 214  
215. 2464  
/note="unlabeled protein product"  
/codon\_start=1  
/protein\_id="CAD22953.1"  
/db\_xref="GI:18618677"  
/translation="MLSIKSGERIVFTFSCQSPENHFVIEIQINDCKMSGPCPGEVQ

LQPSLPLTLNRTFIMVYKAKSGLGLQPSIPRLQIGPESCPDQVTHISGRID  
ATVRIIGTFSNGVSRIRKQEGVMALHLPFHRNVSFISIRARSIRKICLIESV  
FEGESLITLMSANYPEBPEDELMQWPAHLPASVSLFNLNSNCKEEREVEY  
IFGSTDLNEFLIEDKOPNMAFNILLOGDDOASRGLRLOFOVLYVJPONESN  
KLYVVDLNERMMSLTIEPRPKOSRKVFLBERTGSLNLTGSGSHKLSFL  
CDLRLRMNVEKTSCTDHYCORRSLOVPSLILHLPVLIHPSMVLVYKORIS  
LVLPVAKLQOHTHEKPCNTSFYSVAIRBODLYRFSRFGSGISIKOQVONISVT  
LNTFAPSFOEASRQGLTVSFLPYKEGAVFTVDTYSKVLRLPMDRGLPSLTV  
SNMISVPRDQVACLTFPERKSGVQGTARAFMI IQORTRAEIIFSLIEDVLPKSFH  
HHSFWNINSNCSPTSGKDLISVTLTPRTVDLTVIILAAGGVLLLSALGLIIC  
YKKKKKTNNKPAVGIVYNGNINTEMGDSKPRKERTMTGPMQMSSTRTPMVCYR  
IPAPSCSORWPTPTRSRAPMGSVLPHPVPVAPGQLQSWPLRSHLLAPLILVNRIT  
PSPIPTMCM"

BASE COUNT 3'UTR 2465. 5897 2465. 5897 1364 g 1467 t 6 others  
ORIGIN  
Query Match 87.5%; Score 5395; DB 6; Length 5897;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 5865; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

QY 291 CCTGAACCTGCGGGGCTCTATGCACTGCTAGAGGCTTCTGCTGAGTGGCGGCGCT 350  
DB 25 CCTGAACCTGCGGGGCTCTATGCACTGCTAGAGGCTTCTGCTGAGTGGCGGCGCT 84  
QY 351 GCGCGCGGGGCGAGAGCTTTTGAATGCTCTGCGCAGAGAACCAATTACAGTTCT 410  
DB 85 GCGCGCGGGGCGAGAGCTTTTGAATGCTCTGCGCAGAGAACCAATTACAGTTCT 144  
QY 411 CATTAAGCTGGGAGCCCGCACTCTGCTGGCAAAACCGTTTCAATGTCATTTCTTAAAG 470  
DB 145 CATTAAGCTGGGAGCCCGCACTCTGCTGGCAAAACCGTTTCAATGTCATTTCTTAAAG 204  
QY 471 ACATATAACATGTTGTCCATCAAGTCTGAGAAAGAAATAGCTTTACCTTTAGTGCGCA 530  
DB 205 ACATATAACATGTTGTCCATCAAGTCTGAGAAAGAAATAGCTTTACCTTTAGTGCGCA 264  
QY 531 GAGTCTCGAATCACTTTGTATAGAGATCCAGAAAATATTGACTGTATGTCAGGCC 590  
DB 265 GAGTCTCGAATCACTTTGTATAGAGATCCAGAAAATATTGACTGTATGTCAGGCC 324  
QY 591 ATGTCCTTTTGGGAGGTTCACTTCACTGAGCCCTGCAATGCTGTGTCCTACCTCAACAG 650  
DB 325 ATGTCCTTTTGGGAGGTTCACTTCACTGAGCCCTGCAATGCTGTGTCCTACCTCAACAG 384  
QY 651 AACTTTCATCTGGGATGTCAAAGCTCATTAAGACATCGGTTTAAAGTGCAGTTCAT 710  
DB 385 AACTTTCATCTGGGATGTCAAAGCTCATTAAGACATCGGTTTAAAGTGCAGTTCAT 444  
QY 711 CCTGCGCTGAGGCAAGATCGGTCCGGGTGAGAGCTGCCAGAGGATCACTCATCTCAT 770  
DB 445 CCTGCGCTGAGGCAAGATCGGTCCGGGTGAGAGCTGCCAGAGGATCACTCATCTCAT 504  
QY 771 CAGCGCGCGAATTCATGTCACCGGTGTGAGATTCGGAACCTTCTGACGAATGAGCATCT 830  
DB 505 CAGCGCGCGAATTCATGTCACCGGTGTGAGATTCGGAACCTTCTGACGAATGAGCATCT 564  
QY 831 GTCCCGGATCAAGATGCAAGAGAGAGTGAATGAGCTTACCTCCATGTTCCACCC 890  
DB 565 GTCCCGGATCAAGATGCAAGAGAGAGTGAATGAGCTTACCTCCATGTTCCACCC 624  
QY 891 CAGAAATGTCTCGGCTTCAAGATTGCAAAACCGCTCATCTAATAACGCTGTGATCAT 950  
DB 625 CAGAAATGTCTCGGCTTCAAGATTGCAAAACCGCTCATCTAATAACGCTGTGATCAT 684  
QY 951 CAGATCTGTGTGAGGGTGAAGGCTCAGCAACCTGATGTCTGCCAATACCAGAAAG 1010  
DB 685 CAGATCTGTGTGAGGGTGAAGGCTCAGCAACCTGATGTCTGCCAATACCAGAAAG 744  
QY 1011 CTTCCTGAGATGAGTCAATGAGTGGAGTGTGCTGCTGTCAGACACTGCGGGCCAG 1070  
DB 745 CTTCCTGAGATGAGTCAATGAGTGGAGTGTGCTGCTGTCAGACACTGCGGGCCAG 804



QY	1071	CGTCCCTTCCCTCAACTTCAACCTCTCCAACTGTGAGAGGAAGAGACGGGTGTAAATA	1130
Db	805	CGTCTCCCTTCTCAACTTCAACCTCTTCCAACTGTGAGAGGAAGAGACGGGTGTAAATA	864
QY	1131	CTACATCCCGGGGCTCCACCAACCAACCCCGAGGTGTCTCAAGCTGGAGGACAAGACCTGG	1190
Db	865	CTACATCCCGGGGCTCCACCAACCAACCCCGAGGTGTCTCAAGCTGGAGGACAAGACCTGG	924
QY	1191	GAACATGGCGGGGGAATTTCAACTTCTCTGTCGAAGCTGTGACCAAGATGCCCCAAAGTCC	1250
Db	925	GAACATGGCGGGGGAATTTCAACTTCTCTGTCGAAGCTGTGACCAAGATGCCCCAAAGTCC	984
QY	1251	AGGGAATCCTCCGGCTGCGCATCTTCCAAAGTTTGTGTCCAAATATCCACAAATGAAAGCATTA	1310
Db	985	AGGGAATCCTCCGGCTGCGCATCTTCCAAAGTTTGTGTCCAAATATCCACAAATGAAAGCATTA	1044
QY	1311	AATCTACGTGTGTGACTTGAGTAATGACGAGCCATGTCACTCAACCATCGAGCCACGGCC	1370
Db	1045	AATCTACGTGTGTGACTTGAGTAATGACGAGCCATGTCACTCAACCATCGAGCCACGGCC	1104
QY	1371	CGTCAAAACAGACCGCAAGTTTGTTCCTGGGCTGTTTCGTGTGTCTAGAAATCTCGAACCTG	1430
Db	1105	CGTCAAAACAGACCGCAAGTTTGTTCCTGGGCTGTTTCGTGTGTCTAGAAATCTCGAACCTG	1164
QY	1431	CAGTATGCAACTCACCTCGACATCTGGCTCCAAACACAAATCTCTCTTCTTGTATGA	1490
Db	1165	CAGTATGCAACTCACCTCGACATCTGGCTCCAAACACAAATCTCTCTTCTTGTATGA	1224
QY	1491	TCTGACACGTGTGTGATGAATGTGTGAAAAACAATAAGCTGCACAGACACCGGTACTG	1550
Db	1225	TCTGACACGTGTGTGATGAATGTGTGAAAAACAATAAGCTGCACAGACACCGGTACTG	1284
QY	1551	CCAAAGGAATCTCTACTCACTCCAGGTGCCCCAGTGCACATCTCTCAACCTGCTGTGAGCT	1610
Db	1285	CCAAAGGAATCTCTACTCACTCCAGGTGCCCCAGTGCACATCTCTCAACCTGCTGTGAGCT	1344
QY	1611	GCATGACTTCTCTCTGGAAGCTGTGTGTGCCAAGGACAGGCTTCAAGCTGTGTGTGTGCC	1670
Db	1345	GCATGACTTCTCTCTGGAAGCTGTGTGTGCCAAGGACAGGCTTCAAGCTGTGTGTGTGCC	1404
QY	1671	AGCCACAGAGCTGACGACGATACACAGAAAGCCCTGGAAACACAGACTTCAAGTAACT	1730
Db	1405	AGCCACAGAGCTGACGACGATACACAGAAAGCCCTGGAAACACAGACTTCAAGTAACT	1464
QY	1731	CGTGGCCAGTGCATTAACCCACACAGGACCTGTACTTCCGGCTCTTCTTCCGCCGAGAGCTC	1790
Db	1465	CGTGGCCAGTGCATTAACCCACACAGGACCTGTACTTCCGGCTCTTCTTCCGCCGAGAGCTC	1524
QY	1791	TATCAAGCAGATCCAGGTGAAGACAGAAATCTCGGTGACACCTTCTGCACTTTTGCCCCAG	1850
Db	1525	TATCAAGCAGATCCAGGTGAAGACAGAAATCTCGGTGACACCTTCTGCACTTTTGCCCCAG	1584
QY	1851	CTTTCCAACAAAGGCTCTCCAGGACAGGCTCTGACGGTGTCTTTTATTACTTAATTTCAAAAGA	1910
Db	1585	CTTTCCAACAAAGGCTCTCCAGGACAGGCTCTGACGGTGTCTTTTATTACTTAATTTCAAAAGA	1644
QY	1911	GGAAGAGCGTTTTCACGAGTACCCCTGACACAAAAAAGCAAGGCTTACCTGAGGACCCCCCA	1970
Db	1645	GGAAGAGCGTTTTCACGAGTACCCCTGACACAAAAAAGCAAGGCTTACCTGAGGACCCCCCA	1704
QY	1971	CTGGGACCGGGGGCTTGCATCTTCACTCTGTGTCTCTGGAACATCAAGCGTGCACAGAGA	2030
Db	1705	CTGGGACCGGGGGCTTGCATCTTCACTCTGTGTCTCTGGAACATCAAGCGTGCACAGAGA	1764
QY	2031	CCAGGTGGCTTGTCTGACTTTCTTTTAAAGACGGACGGGCTGTGCTGACCAAGACGGGCG	2090
Db	1765	CCAGGTGGCTTGTCTGACTTTCTTTTAAAGACGGACGGGCTGTGCTGACCAAGACGGGCG	1824
QY	2091	CGCAATTCAATGATTCACGAGACAGCGGACCCGGGCTGAGAGATCTTCAAGCTGTGAGAGA	2150
Db	1825	CGCAATTCAATGATTCACGAGACAGCGGACCCGGGCTGAGAGATCTTCAAGCTGTGAGAGA	1884
QY	2151	GGATGTGCTCCCAAGCCAGACTTTCACATCAAGCTTCTGGGTCAACATCTCTAATCTG	2210

Db	1885	GGATGTGCTCCCAAGCCAAAGCTTCCACATCACAGCTTCTGGGTCAACATCTCTAATCTG	1944
OY	2211	CAGCCCAACGAGCGGCAAGAGAGCTAGACCTGCTCTTCTCGGTGACACTTACCCCAAGAC	2270
Db	1945	CAGCCCAACGAGCGGCAAGAGAGCTAGACCTGCTCTTCTCGGTGACACTTACCCCAAGAC	2004
OY	2271	TGTGGACTTGACTGTCAATCTCCATCCGACCGGTGGAGGTGAGTCTTACTGCTGTCCG	2330
Db	2005	TGTGGACTTGACTGTCAATCTCCATCCGACCGGTGGAGGTGAGTCTTACTGCTGTCCG	2064
OY	2231	CTTCGGGCTCACTCAATTTGTGTGTGAAAAAGAAAGAAAAAGAACAAAGAGGCCCGC	2390
Db	2065	CTTCGGGCTCACTCAATTTGTGTGTGAAAAAGAAAGAAAAAGAACAAAGAGGCCCGC	2124
OY	2391	TGTGGGTATCTACCAATGGCAACATCAATTACTAGATGCCAGAGGCCCAAAAGTTTCA	2450
Db	2125	TGTGGGTATCTACCAATGGCAACATCAATTACTAGATGCC-AGGACGCCAAAAGTTTCA	2183
OY	2451	GAAGGGCGAAGGACAATGACTCCATGTGTATGACGTCAATCGAGACACCATGTATTA	2510
Db	2184	GAAGGGCGAAGGACAATGACTCCATGTGTATGACGTCAATCGAGAGACCATGTATTA	2243
OY	2511	TGGGCATCTGTANAGGATTCACGGGCTCTTCCTGACACCGAAGGTGACACCTAACG	2570
Db	2244	TGGGCATCTGTANAGGATTCACGGGCTCTTCCTGACACCGAAGGTGACACCTAACG	2303
OY	2571	GCCTTTCAGGGGCAACATGGGGGCTGTCTCTCCCTCCCAACCCCAATATGCTCAAGGC	2630
Db	2304	GCCTTTCAGGGGCAACATGGGGGCTGTCTCTCCCTCCCAACCCCAATATGCTCAAGGC	2363
OY	2631	CCCAACTGCAAAAGTTGGCCACTGAGAGCCACTCTCGCTCCCTCTGTAGTGTAGAG	2690
Db	2364	CCCAACTGCAAAAGTTGGCCACTGAGAGCCACTCTCGCTCCCTCTGTAGTGTAGAG	2423
OY	2691	TGAACCGTACACCTTCTCCCATCCCAACAATGGGGATGTAAAGCAGCAAGACACAT	2750
Db	2424	TGAACCGTACACCTTCTCCCATCCCAACAATGGGGATGTAAAGCAGCAAGACACAT	2483
OY	2751	TCCCTTACTGAACACTCAGAGAGCCACTGAGAGCCAGAGAAATTAATTATCATTTCCAGAC	2810
Db	2484	TCCCTTACTGAACACTCAGAGAGCCACTGAGAGCCAGAGAAATTAATTATCATTTCCAGAC	2543
OY	2811	GCTTTGCTGAATTCATTAAGCAGGGCACTGAGACACCCGTCGCTGTTCTTAACAGAAA	2870
Db	2544	GCTTTGCTGAATTCATTAAGCAGGGCACTGAGACACCCGTCGCTGTTCTTAACAGAAA	2603
OY	2871	TCTTAAAGAAAGGAATTAATACAGAGAAACAGCAGAGGTTTTCTGAGACCGCCAAC	2930
Db	2604	TCTTAAAGAAAGGAATTAATACAGAGAAACAGCAGAGGTTTTCTGAGACCGCCAAC	2663
OY	2931	TTTCACTTTGCTCAGTGAACCTCATTTCTAAGGGCAAGCAATTTGAAAATGATAAATTCACATC	2990
Db	2664	TTTCACTTTGCTCAGTGAACCTCATTTCTAAGGGCAAGCAATTTGAAAATGATAAATTCACATC	2723
OY	2991	TGATATACGTCATGACAGCTCATGTGCTCTCAACTTAAGGCTGTGCGGTTAAGCCAGCTG	3050
Db	2724	TGATATACGTCATGACAGCTCATGTGCTCTCAACTTAAGGCTGTGCGGTTAAGCCAGCTG	2783
OY	3051	TAAATGAGAGAGAGGCTGTAGTCACTTACATAGGATAGGTTGACAGAACCCCTGGATTCAG	3110
Db	2784	TAAATGAGAGAGAGGCTGTAGTCACTTACATAGGATAGGTTGACAGAACCCCTGGATTCAG	2843
OY	3111	AGTGTAAACAGAGGCTTGCCCTTTCAGAGCAACAATTCCAATTCAGAGAGCTTACT	3170
Db	2844	AGTGTAAACAGAGGCTTGCCCTTTCAGAGCAACAATTCCAATTCAGAGAGCTTACT	2903
OY	3171	GAGGTCCCTACTCTCACTGGGGTCCCAAGATGAAACGACAAATGTGCTTTTATTAATTT	3230
Db	2904	GAGGTCCCTACTCTCACTGGGGTCCCAAGATGAAACGACAAATGTGCTTTTATTAATTT	2963
OY	3231	ATTATATTGGTGTCTGTGTATTTAAGATCAAAATGTATACACCTAGCTCTTTTC	3290

Db 2964 ATTATTTGGTGTCTCTGTCTTATTTAAGAGATCAAAATGTATACCACTAGCTTTTTC 3023  
Qy 3291 ACCTGACTTATATTAATCATCTAATCTGTTGGATGCTGGTTGACTTCTACTG 3350  
Db 3024 ACCTGACTTATATTAATCATCTAATCTGTTGGATGCTGGTTGACTTCTACTG 3083  
Qy 3351 ACCGCTAGATTAACGTGTGCTGTCTCCCGAGTGTGGGAATATTTAATCTGTCCA 3410  
Db 3084 ACCGCTAGATTAACGTGTGCTGTCTCCCGAGTGTGGGAATATTTAATCTGTCCA 3143  
Qy 3411 CCGAAGAAAGATGTGTGTGTGAGAGCATTGACATATCTGCTTTGATAGAGACTT 3470  
Db 3144 CCGAAGAAAGATGTGTGTGTGAGAGCATTGACATATCTGCTTTGATAGAGACTT 3203  
Qy 3471 CCTGATTTCTTAGGTGCTGTGTGTATCCATTTGGAATTTCAATCTGTAATCCCAT 3530  
Db 3204 CCTGATTTCTTAGGTGCTGTGTGTATCCATTTGGAATTTCAATCTGTAATCCCAT 3263  
Qy 3531 GTCTCTATGTCTTAGCATTAAGAAATTTCTCAAGTTTCCATGTGCGGTTCTCTAGC 3590  
Db 3264 GTCTCTATGTCTTAGCATTAAGAAATTTCTCAAGTTTCCATGTGCGGTTCTCTAGC 3323  
Qy 3591 TGGAGCAATCTTGAATTTAAGAAATTTAAGAAATTTCTCATCTCTAATAAATG 3650  
Db 3324 TGGAGCAATCTTGAATTTAAGAAATTTAAGAAATTTCTCATCTCTAATAAATG 3383  
Qy 3651 TTTAAATATATACCAACAGTGGCCCCCTGATTAATTTCTGTTGCCACTGCAACCAT 3710  
Db 3384 TTTAAATATATACCAACAGTGGCCCCCTGATTAATTTCTGTTGCCACTGCAACCAT 3443  
Qy 3711 TACTGTAGCTTTAAACACACATTAAGTTAAGTCTGGGATCAGATTTCCAAAT 3770  
Db 3444 TACTGTAGCTTTAAACACACATTAAGTTAAGTCTGGGATCAGATTTCCAAAT 3503  
Qy 3771 GATGTCCCTGAATGAATCAAGGTGTGAGAGAGTGTGCTCTTGAAGGCTCTAG 3830  
Db 3504 GATGTCCCTGAATGAATCAAGGTGTGAGAGAGTGTGCTCTTGAAGGCTCTAG 3563  
Qy 3831 GGAGAACCGGTTCTTGGCCATTTCAAGCTTTAGAGGCTGCTGATTTCCAGGCTCA 3890  
Db 3564 GGAGAACCGGTTCTTGGCCATTTCAAGCTTTAGAGGCTGCTGATTTCCAGGCTCA 3623  
Qy 3891 GTGGCTGTGACCTTTTCTCAATGAGCATCACTGTGACATGCGCTTCCCACTTCCCTC 3950  
Db 3624 GTGGCTGTGACCTTTTCTCAATGAGCATCACTGTGACATGCGCTTCCCACTTCCCTC 3683  
Qy 3951 TTTGACTTCAAGGCCCAACAGGAATCCAGGAATATCTCTCATCTAAGATCCCTTA 4010  
Db 3684 TTTGACTTCAAGGCCCAACAGGAATCCAGGAATATCTCTCATCTAAGATCCCTTA 3743  
Qy 4011 TCATCTCGAAGAGCTTTTGGCATGCAAGACATAGCCACAGGTGGGATTTAGAGCC 4070  
Db 3744 TCATCTCGAAGAGCTTTTGGCATGCAAGACATAGCCACAGGTGGGATTTAGAGCC 3803  
Qy 4071 AGGACATCTTTGGGGTGTGTTATTTCTGCTTACCACTTCTCTGCACTGACTCCACA 4130  
Db 3804 AGGACATCTTTGGGGTGTGTTATTTCTGCTTACCACTTCTCTGCACTGACTCCACA 3863  
Qy 4131 GGAGAGGCTTCAAAATGATCTGGGCAACAGGAGTGTGTTTGAAGCTTGGGACTTAAC 4190  
Db 3864 GGAGAGGCTTCAAAATGATCTGGGCAACAGGAGTGTGTTTGAAGCTTGGGACTTAAC 3923  
Qy 4191 ACTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4250  
Db 3924 ACTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3983  
Qy 4251 AACACTGTGCTGAGCTGGGACCAAGTCTGCTTTAAGAGGGGTGTCACTTCCACAG 4310  
Db 3984 AACACTGTGCTGAGCTGGGACCAAGTCTGCTTTAAGAGGGGTGTCACTTCCACAG 4043  
Qy 4311 TCACCAAGGCCACATAGGCTTATCACTTCCACAAATAGAGCTTAAGTGTGTTCTA 4370  
Db 4044 TCACCAAGGCCACATAGGCTTATCACTTCCACAAATAGAGCTTAAGTGTGTTCTA 4103

Qy 4371 CTGATCAATGCCCTGACAGGTTGATTTATTTGTAATGAAGAAAGACTGGATTAATC 4430  
Db 4104 CTGATCAATGCCCTGACAGGTTGATTTATTTGTAATGAAGAAAGACTGGATTAATC 4163  
Qy 4431 TCTATACAGTGAATGACATAGAACCAATGTGTCTCACTTACCTTTTCTTTT 4490  
Db 4164 TCTATACAGTGAATGACATAGAACCAATGTGTCTCACTTACCTTTTCTTTT 4223  
Qy 4491 TCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 4550  
Db 4224 TCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 4283  
Qy 4551 TGAAGTGAATGGGCAATCTGAGTCACTGCACTGCTGCTGCTGCTGCTGCTGCTGCT 4610  
Db 4284 TGAAGTGAATGGGCAATCTGAGTCACTGCACTGCTGCTGCTGCTGCTGCTGCTGCT 4343  
Qy 4611 CTCCCACTGAGCTTCCCAAAATGCTGGATCACTGCACTGCTGCTGCTGCTGCTGCT 4670  
Db 4344 CTCCCACTGAGCTTCCCAAAATGCTGGATCACTGCACTGCTGCTGCTGCTGCTGCT 4403  
Qy 4671 TTTTGAATTTTGTAGAGACAGGTTTCAACATGTGGCCAGGCTGTCTCAACTCT 4730  
Db 4404 TTTTGAATTTTGTAGAGACAGGTTTCAACATGTGGCCAGGCTGTCTCAACTCT 4463  
Qy 4731 GGGCTCAAGCAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4790  
Db 4464 GGGCTCAAGCAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4523  
Qy 4791 CCGCATCCAGCCCCACACCTTATTTATACCAATTAACCTGCCATTAATGTGACTTT 4850  
Db 4524 CCGCATCCAGCCCCACACCTTATTTATACCAATTAACCTGCCATTAATGTGACTTT 4583  
Qy 4851 GCTTCTCAACCTCTCTGATCTGAAAGAGAGGATTAATTTAGCTTGTGACACA 4910  
Db 4584 GCTTCTCAACCTCTCTGATCTGAAAGAGAGGATTAATTTAGCTTGTGACACA 4643  
Qy 4911 GTCCCAAGTTCAATTTCTGCGGCAAAAATCTTCTCAAAAATTAATGTAATTTG 4970  
Db 4644 GTCCCAAGTTCAATTTCTGCGGCAAAAATCTTCTCAAAAATTAATGTAATTTG 4703  
Qy 4971 TATTTCAATGAATTTCACTTTGGAATGACACCGCTCAACTTTGTTCAATGGAATGA 5030  
Db 4704 TATTTCAATGAATTTCACTTTGGAATGACACCGCTCAACTTTGTTCAATGGAATGA 4763  
Qy 5031 AGGAATTTTATAGTCTCTTAATGAGGCTGTAAGTGAAGCTCTTGAACATTTCCAGAG 5090  
Db 4764 AGGAATTTTATAGTCTCTTAATGAGGCTGTAAGTGAAGCTCTTGAACATTTCCAGAG 4823  
Qy 5091 GATAGGATATTTTAAATGATGCTTGGCGTTGCTTATGAGCACTTTCCCTTCTGAAGTC 5150  
Db 4824 GATAGGATATTTTAAATGATGCTTGGCGTTGCTTATGAGCACTTTCCCTTCTGAAGTC 4883  
Qy 5151 TGGTTCTGCGCCAGAGACCTTGGCTTGTGAGCCGAGATGTGACCTTGAATTAAGGGC 5210  
Db 4884 TGGTTCTGCGCCAGAGACCTTGGCTTGTGAGCCGAGATGTGACCTTGAATTAAGGGC 4943  
Qy 5211 CAAAGAGGGCTGCGGCTTCTTCCCTCACTAAGAGCCCTTATTTGAATTCAGTGTG 5270  
Db 4944 CAAAGAGGGCTGCGGCTTCTTCCCTCACTAAGAGCCCTTATTTGAATTCAGTGTG 5003  
Qy 5271 GAGCCCTAGCCCTCAATTTCTGACATTTCCCAACCTTCCAGCCCTTCCAGACGACTA 5330  
Db 5004 GAGCCCTAGCCCTCAATTTCTGACATTTCCCAACCTTCCAGCCCTTCCAGACGACTA 5063  
Qy 5331 GGTGCTCTGATTTCCACCAAGTGGATTTGGCTTCTTATAGCTGCTACTTTGTACCA 5390  
Db 5064 GGTGCTCTGATTTCCACCAAGTGGATTTGGCTTCTTATAGCTGCTACTTTGTACCA 5123  
Qy 5391 TCACCAAGCTCACTGTGCTGCAAGACACCAAGTGGCACTTTTCTTCACTGAGGAC 5450  
Db 5124 TCACCAAGCTCACTGTGCTGCAAGACACCAAGTGGCACTTTTCTTCACTGAGGAC 5183

QY	5451	TCAAAACCTCCGAGCAAACTGTGCGCTCCTAGACCAATATTTCTCGAAGCTGAGCCCTCA	5510
Db	5184	TCAAAACCTCCGAGCAAACTGTGCGCTCCTAGACCAATATTTCTCGAAGCTGAGCCCTCA	5243
QY	5511	GTGAAGGGGGCCAGCTGAGGAACCCCTGGCTCTTTTCTTTAAAGCCAGGCCCACTTAC	5570
Db	5244	GTGAAGGGGGCCAGCTGAGGAACCCCTGGCTCTTTTCTTTAAAGCCAGGCCCACTTAC	5303
QY	5571	ATMAAACATTTCCAGGGTCACTGGAACAGTGAAGTGCATTTTGTGAAGCTTACATGCATG	5630
Db	5304	ATMAAACATTTCCAGGGTCACTGGAACAGTGAAGTGCATTTGTGGAAGCCCTACTGATAGT	5363
QY	5631	CCAGGCCCACTGCTCATCCACGGGGCTGCGCAATGCCCTACAGAGGAAGGCCAGGCATGCAGG	5690
Db	5364	CCAGGCCCACTGCTCATCCACGGGGCTGCGCAATGCCCTACAGAGGAAGGCCAGGCATGCAGG	5423
QY	5691	ACTGGTCTCTAATGCTGTGTGATCATTTGCAAGAGGGAAGGCTCTCAAGGAAGTCAACT	5750
Db	5424	ANTGGTCTCTAATGNTGTGTGTGATCATTTGCAAGAGGAAGGCTCTCAAGGAAGTCAACT	5483
QY	5751	GGGACAGACCAAGCCCAACCGGACATGGCCCTTGGTAAAGTTAGCAGACTGGTGTGTGTG	5810
Db	5484	GGGACAGACCAAGCCCAACCGGACATGGCCCTTGGTAAAGTTAGCAGACTGGTGTGTGTG	5543
QY	5811	GATCGCAGTGTCTCATCTGGAATTAATTAATTCATGAGATPACTTTTATAGTGGCAATT	5870
Db	5544	GATCGCAGTGTCTCATCTGGAATTAATTAATTCATGAGATPACTTTTATAGTGGCAATT	5603
QY	5871	TATTCATTTCTGTGCTTTAATTAATTAACAATGTACCAAAAAACAAGTATCAAGCTGTTA	5930
Db	5604	TATTCATTTCTGTGCTTTAATTAATTAACAATGTACCAAAAAACAAGTATCAAGCTGTTA	5663
QY	5931	AGTGCCTTCGAGTACTGTGTCCCTGTGGTTCAGTAGAGGCCCGGTTTCCAGTGTGTGACTG	5990
Db	5664	AGTGCCTTCGAGTACTGTGTCCCTGTGGTTCAGTAGAGGCCCGGTTTCCAGTGTGTGACTG	5723
QY	5991	TGACACAGGCTCAGCATGGGCTCAGCAGATGCTGTCTTAATTTTGTGATATATACAGAAACC	6050
Db	5724	TGACACAGGCTCAGCATGGGCTCAGCAGATGCTGTCTTAATTTTGTGATATATACAGAAACC	5783
QY	6051	AGGCTTTGGGATACAAAGTTCTTTTCCCTCTTCATTTGATGCGCGACCTGTGGAAGCAGAT	6110
Db	5784	AGGCTTTGGGATACAAAGTTCTTTTCCCTCTTCATTTGATGCGCGACCTGTGGAAGCAGAT	5843
QY	6111	GTTTTGTTCGCGGAATATAAAATAATAGTCTTGAAGTCTCGCCAAAAAATAAAAA 6163	
Db	5844	GTTTTGTTCGCGGAATATAAAATAATAGTCTTGAAGTCTCGCCAAAAAATAAAAA 5896	
RESULT 4			
AF468010			
LOCUS	AF468010	5978 bp	mRNA
DEFINITION	Homo sapiens NCSE135 mRNA, complete cds.	linear	PRI 02-FEB-2003
ACCESSION	AF468010		
VERSION	AF468010.1	GI:28194091	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 5978)		
AUTHORS	Hooper,J.D., Ziljstra,A., Aimes,R.T., Claassen,G.F. and Ougley,J.P.		
TITLE	Subtractive Immunization Identifies NCSE-135, a Novel Cell Surface Glycoprotein		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 5978)		
AUTHORS	Hooper,J.D., Ziljstra,A., Aimes,R.T., Claassen,G.F. and Ougley,J.P.		
TITLE	Direct Submision		
JOURNAL	Submitted (14-JAN-2002) Cell Biology, Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA		
FEATURES	Location/Qualifiers		

Source	1. .5978	/organism="Homo sapiens"	/mol_type="mRNA"	/db_xref="taxon:9606"	/chromosome="3"	/map="3p21.32"	111. .2621	/note="novel cell surface glycoprotein 135 kDa"	/codon_start=1	/product="NCSG135"	/protein_id="AA033397.1"	/db_xref="GI:28194092"	/translation="MAGINGVSGIALIGVLLIGARLPREGAAPEIPLPRESNTYI.IKGLPTPLIAPCTIVISKRHITMLSTSGRIYPTPSCGSPENHFVETIQNDMSGPPCEGEVQLQPSTLSLEPLTRTFIWDYRAKHSIGLEIQSFIPRIQIPEBSCDGVATLHISGRIDATVIRIGTCSNGTWSRIIMQGVKQALPLPMFHENWGSFIIANRSIKRLCIESIVPEFGESATLMSANYPEGPEPDELMTMQEVPYPAHLBASVFILPNFINSCEBKEEVEYIIPGSTTNPPEFLDKDPQNMAGNPLSLQGGCDADQSGILRILOFQVILVLOHDPENSKITVYNDLSNEBMSLITLPRPVKSRKPYGCFCLSEPTCSNLTITSGSKHISFLCDLITRLMKNVEKTTISCTDHRICQKRSYSLOQPSDILHLPVLDHPSMKLIVPDELSTLVLVPAQKIQOHTHEKPCNTSFYSVLVSAIIPQDILYFSGFCGSGIKQIQVKONI SVLTRKPAFSPROBASQGLTVSFLPYKEEVEVPIVPTDYSKVILRTPMNBORGLBSTISVWNI SVPRDQVACLTFEKRSGVVCOTGAFLMIIGORAFIEI FSLDEDVLPRSPFHHSFMWNITNSGSPGSKDILFSLTLPRTYDILVILAAVGGVLLISALGLIICVCYKKKKKTKNKPANVEIYNQINITMBPRQKROKRNKDNHIVAVIETPLMVTGHILQDSSGSFLQPEVDITRPFQGTMSVCPSPEPTISRAPTALDSEPEPPSPPESESEPTYSHPNNGVSSKDDIPLLSIQEBWEPAE"
Query Match	81.1%	Score 4996;	DB 9;	Length 5978;									
Best Local Similarity	99.7%	Pred. No. 0;											
Matches 5966;	Conservative	0;	Mismatches 11;	Indels 4;	Gaps 3;								
BASE COUNT	1457 a	1637 c	1407 g	1477 t									
ORIGIN													
Query	173	GC	GC	GC	GC	GC	GC	GC	GC	GC	GC	GC	GC
Db	1	GC	GC	GC	GC	GC	GC	GC	GC	GC	GC	GC	GC
Query	233	G	C	C	C	C	C	C	C	C	C	C	C
Db	61	G	C	C	C	C	C	C	C	C	C	C	C
Query	293	T	G	A	A	C	T	G	G	G	G	T	T
Db	121	T	G	A	A	C	T	G	G	G	G	T	T
Query	353	C	G	C	G	C	G	G	G	G	G	C	A
Db	181	C	G	C	G	C	G	G	G	G	G	C	A
Query	413	T	A	A	A	C	T	G	G	G	G	T	T
Db	241	T	A	A	A	C	T	G	G	G	G	T	T
Query	473	A	T	A	T	A	C	A	T	G	T	G	C
Db	301	A	T	A	T	A	C	A	T	G	T	G	C
Query	533	G	T	C	T	G	A	G	A	T	C	A	T
Db	361	G	T	C	T	G	A	G	A	T	C	A	T
Query	593	G	T	C	T	T	T	G	G	A	G	T	T
Db	421	G	T	C	T	T	T	G	G	A	G	T	T
Query	653	C	T	T	T	C	A	T	C	T	G	A	T
Db	481	C	T	T	T	C	A	T	C	T	G	A	T
Query	713	C	T	G	C	C	T	G	A	G	C	A	T
Db	541	C	T	G	C	C	T	G	A	G	C	A	T

QY 773 GCGGCGGAATGATGCCACCGTGTGAGATGCGAACTTTGACGAATGACACTGTGT 832  
DB 601 GCGGCGGAATGATGCCACCGTGTGAGATGCGAACTTTGACGAATGACACTGTGT 660  
QY 833 CCGGGATTCAGAGATGCAAGAGAGATGAAAAATGACCTTACCTCCCATGATGTTACACCCA 892  
DB 661 CCGGGATTCAGAGATGCAAGAGAGATGAAAAATGACCTTACCTCCCATGATGTTACACCCA 720  
QY 893 GAAATGCTTCGGGCTTCAGACTTGCAAAACCGCTCATCTATATAAAGCTGTGATCATCG 952  
DB 721 GAAATGCTTCGGGCTTCAGACTTGCAAAACCGCTCATCTATATAAAGCTGTGATCATCG 780  
QY 953 AGTGTGTGTTGAGGGTGAAGGCTCAGCAACCTGTATGTCTGCAACTACCAAGAAAGCT 1012  
DB 781 AGTGTGTGTTGAGGGTGAAGGCTCAGCAACCTGTATGTCTGCAACTACCAAGAAAGCT 840  
QY 1013 TCCCTGAGAGATGAGCTCATGACCTGCGACTTTGTCTGTCACACCTGCGGCGCAGAG 1072  
DB 841 TCCCTGAGAGATGAGCTCATGACCTGCGACTTTGTCTGTCACACCTGCGGCGCAGAG 900  
QY 1073 TCTCTTCTCCTCAACTTCAACCTCTCCCACTGTGAGAGAGAGAGAGCGGGTTGAATACT 1132  
DB 901 TCTCTTCTCCTCAACTTCAACCTCTCTCACTGTGAGAGAGAGAGAGCGGGTTGAATACT 960  
QY 1133 ACATCCCGGGCTCCACCAACCCCGAGGTGTTCAAGCTGAGAGCAAGCAGCTGGGA 1192  
DB 961 ACATCCCGGGCTCCACCAACCCCGAGGTGTTCAAGCTGAGAGCAAGCAGCTGGGA 1020  
QY 1193 ACATGCGGGGAACTTCAACCTCTCTCTGCAAGCTGTGACCAAGATGCCCAAAATGCCAG 1252  
DB 1021 ACATGCGGGGAACTTCAACCTCTCTCTGCAAGCTGTGACCAAGATGCCCAAAATGCCAG 1080  
QY 1253 GGAATCCCGGGCTGACAGTTTCAAGTTTGTGTCACATCCCAAAATGAAAGCAATAAA 1312  
DB 1081 GGAATCCCGGGCTGACAGTTTCAAGTTTGTGTCACATCCCAAAATGAAAGCAATAAA 1140  
QY 1313 TCTAGTGTGTTGACTTGAATGAGAGAGCACTGTCACATCCAGCCAGCCGCGCG 1372  
DB 1141 TCTAGTGTGTTGACTTGAATGAGAGAGCACTGTCACATCCAGCCAGCCGCGCG 1200  
QY 1373 TCAAAAG 1432  
DB 1141 TCAAAAG 1200  
QY 1433 GTAGCAACCTCAGCTGACATCTGAGCTCAAAACAATAATCTCTTCTGTGATGATC 1492  
DB 1261 GTAGCAACCTCAGCTGACATCTGAGCTCAAAACAATAATCTCTTCTGTGATGATC 1320  
QY 1493 TGAACGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1552  
DB 1321 TGAACGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1380  
QY 1553 AAAGGAATCTTACTCACTCAGAGTGCAGAGTGCATCTTCCAGCTGTGAGAGTGC 1612  
DB 1381 AAAGGAATCTTACTCACTCAGAGTGCAGAGTGCATCTTCCAGCTGTGAGAGTGC 1440  
QY 1613 ATGACTTCTCTGGAAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1672  
DB 1441 ATGACTTCTCTGGAAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1500  
QY 1673 CCGGAGAGCTGAGCAGCAGATACACGAGAAAGCCCTGCAACACAGCTTCACTACCTCG 1732  
DB 1501 CCGGAGAGCTGAGCAGCAGATACACGAGAAAGCCCTGCAACACAGCTTCACTACCTCG 1560  
QY 1733 TGGCAGATGCTATACCAAGCAGAGACTGTATCTTGGCTCTTGTGCGCGGAGGCTCTA 1792  
DB 1561 TGGCAGATGCTATACCAAGCAGAGACTGTATCTTGGCTCTTGTGCGCGGAGGCTCTA 1620  
QY 1793 TCAAGCAGATCAGGTGAGCAGAACTCTCGGTGACCTTGGCACTTTGGCCCGAGCT 1852  
DB 1621 TCAAGCAGATCAGGTGAGCAGAACTCTCGGTGACCTTGGCACTTTGGCCCGAGCT 1680

QY 1853 TCCAAAGAGAGCTTCAGAGAGAGCTGACGCTGCTCTTATACCTTATTTCAAAGAG 1912  
DB 1681 TCCGAAGAGAGCTTCAGAGAGAGCTGACGCTGCTCTTATACCTTATTTCAAAGAG 1740  
QY 1913 AAGGGTTTTCAAGGTGACCCCTGACACAAAAGAGAGCTACCTGAGAGAGAGAGAGAG 1972  
DB 1741 AAGGGTTTTCAAGGTGACCCCTGACACAAAAGAGAGAGCTACCTGAGAGAGAGAGAG 1800  
QY 1973 GGGAGCGGGGCTGACATCCCTCAGCTGTGTCTCTGCAACATCAGCAGCTGCCAGAGAG 2032  
DB 1801 GGGAGCGGGGCTGACATCCCTCAGCTGTGTCTCTGCAACATCAGTGTGCCAGAGAGAG 1860  
QY 2033 AGTGGCTGCTGACTTCTTTAAGAGCGGAGCGGCTGTGCTGCGACAGAGCGCGG 2092  
DB 1861 AGTGGCTGCTGACTTCTTTAAGAGCGGAGCGGCTGTGCTGCGACAGAGCGCGG 1920  
QY 2093 CATTCATGATCATTCAG 2152  
DB 1921 CATTCATGATCATTCAG 1980  
QY 2153 ATGTGCTCCCAAGCCAGCTTCACATCAGAGCTTCTGGGTCAACATCTCTAACTGCA 2212  
DB 1981 ATGTGCTCCCAAGCCAGCTTCACATCAGAGCTTCTGGGTCAACATCTCTAACTGCA 2040  
QY 2213 GCCCAAGAGCGGCAAGCAGCTGACCTGCTTCTCGGTGACATTTACCAGAGAGCTG 2272  
DB 2041 GCCCAAGAGCGGCAAGCAGCTGACCTGCTTCTCGGTGACATTTACCAGAGAGCTG 2100  
QY 2273 TGGACTGATGATCATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2332  
DB 2101 TGGACTGATGATCATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2160  
QY 2333 TCGGGCTCATCTTGTGTGTGAAAAAGAGAAAAAGAGCAAAAGAGAGAGAGAGAGAG 2392  
DB 2161 TCGGGCTCATCTTGTGTGTGAAAAAGAGAAAAAGAGCAAAAGAGAGAGAGAGAGAG 2220  
QY 2393 TGGATATCTTAATGAGCAATCAATCTGAGATGCCAGGCAAGCAAAAGATTGAGA 2452  
DB 2221 TGGATATCTTAATGAGCAATCAATCTGAGATGCCAGGCAAGCAAAAGATTGAGA 2280  
QY 2453 AAGGGGAAAGAGCAATGATCCCATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2512  
DB 2281 AAGGGGAAAGAGCAATGATCCCATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2340  
QY 2513 GGCATCTGCTAAGAGATTCAGAGAGCTCTCTCTGAGCAGAGAGTGAACACTACCGG 2572  
DB 2341 GGCATCTGCTAAGAGATTCAGAGAGCTCTCTCTGAGCAGAGAGTGAACACTACCGG 2400  
QY 2573 CGTTCAGGGCAATGAGGAGTCTGCTCTCTCCCAACCAACATATGCTCCAGGGCC 2632  
DB 2401 CGTTCAGGGCAATGAGGAGTCTGCTCTCTCCCAACCAACATATGCTCCAGGGCC 2460  
QY 2633 CAATGCAAAAGTTGGCACTGAGAGCACTCTGCTGCTCTCTCTGAGTCTGAGAGTG 2692  
DB 2461 CAATGCAAAAGTTGGCACTGAGAGCACTCTGCTGCTCTCTCTGAGTCTGAGAGTG 2520  
QY 2693 AACCGTACACTTCTCCCAATGAGGAGTGAAGCAGAGAGAGAGAGAGAGAGAGAGAG 2752  
DB 2521 AACCGTACACTTCTCCCAATGAGGAGTGAAGCAGAGAGAGAGAGAGAGAGAGAGAG 2580  
QY 2753 CCTTACTGAGCACTGAG 2812  
DB 2581 CCTTACTGAGCACTGAG 2640  
QY 2813 TTTGCTGATTTCAATGAG 2872  
DB 2641 TTTGCTGATTTCAATGAG 2700  
QY 2873 CTAAAGAGAGAGATTTATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2932  
DB 2701 CTAAAGAGAGAGATTTATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2760  
QY 2933 CAATTTGCTCACTGAGATCTATTTTAAAGGCAAGAGAGAGAGAGAGAGAGAGAGAG 2992

```
Db 2761 CACATTGCTCAGGAGCTCATTTAAGGCGAAGCAATTGAAATGATGAAATTCATCTCG 2820
Qy 2993 GATACAGTCATGACAGCTCATGTCCTCTCAACTTGAAGGCTGCGGTTAGCCAGCCTGA 3052
Db 2821 GATACAGTCATGACAGCTCATGTCCTCTCAACTTGAAGGCTGCGGTTAGCCAGCCTGA 2880
Qy 3053 ATGAGAGAGAGAGGCTGAGTCAGCTAGCATGAGGTTGACAGCAAGCCCTGGAATTCAGAG 3112
Db 2881 ATGAGAGAGAGAGGCTGAGTCAGCTAGCATGAGGTTGACAGCAAGCCCTGGAATTCAGAG 2940
Qy 3113 TGTAAACAGAGGCTTGCCTCTTCAAGACACAGTTCCAAATTCAGAGAGCCTTCTGA 3172
Db 2941 TGTAAACAGAGGCTTGCCTCTTCAAGACACAGTTCCAAATTCAGAGAGCCTTCTGA 3000
Qy 3173 GGTCCCTACTCTCAGTGGGGTCCCGAGATGAAAGCAATGTCCTTTTATTTAT 3232
Db 3001 GGTCCCTACTCTCAGTGGGGTCCCGAGATGAAAGCAATGTCCTTTTATTTAT 3060
Qy 3233 TTATTTGTCGTCCTGCTGTTATTTAAGATCAATGATTAACCACTAGCTCTTTTCA 3292
Db 3061 TTATTTGTCGTCCTGCTGTTATTTAAGATCAATGATTAACCACTAGCTCTTTTCA 3120
Qy 3293 CTGACTTAACTAATCTCATACTAATCTGTTGGATGCTGGGTTGACTTCTACTGAC 3352
Db 3121 CTGACTTAACTAATCTCATACTAATCTGTTGGATGCTGGGTTGACTTCTACTGAC 3180
Qy 3353 CGCTAGATAAGCGTGCGTCCCGCAGTGCGGGAATTAATTTACATCTGTCACAC 3412
Db 3181 CGCTAGATAAGCGTGCGTCCCGCAGTGCGGGAATTAATTTACATCTGTCACAC 3240
Qy 3413 AGAAAAAGATGTCGTCGTTGAGCAGCATTTGACACATCTGCTTGAATAGAGCTTCC 3472
Db 3241 AGAAAAAGATGTCGTCGTTGAGCAGCATTTGACACATCTGCTTGAATAGAGCTTCC 3300
Qy 3473 TGAATCTCTAGGTCGCTGCTGCTGATTCCTATGTCGAAATTCATCTGATTCCTATGT 3532
Db 3301 TGAATCTCTAGGTCGCTGCTGCTGATTCCTATGTCGAAATTCATCTGATTCCTATGT 3360
Qy 3533 CCTATGTCCTGAGCAATTAAGAAATTTCTCAAGTTTCAATGTCGCTTCTCTAGCTG 3592
Db 3361 CCTATGTCCTGAGCAATTAAGAAATTTCTCAAGTTTCAATGTCGCTTCTCTAGCTG 3420
Qy 3593 CAGCAATCTTTGACATTTAAAGAGAAATTTAGAGATATTCATCTCTCTAAATAATGT 3652
Db 3421 CAGCAATCTTTGACATTTAAAGAGAAATTTAGAGATATTCATCTCTCTAAATAATGT 3480
Qy 3653 TAAATATATACCAACAGTGGCCCTGATTAAGTTTCTGTTGCCATGCAACCATTA 3712
Db 3481 TAAATATATACCAACAGTGGCCCTGATTAAGTTTCTGTTGCCATGCAACCATTA 3540
Qy 3713 CTGTCGATGTTAAAAACAACATTTAGCTTATAGTCTGGGATTCGAATTTCCAAATG 3772
Db 3541 CTGTCGATGTTAAAAACAACATTTAGCTTATAGTCTGGGATTCGAATTTCCAAATG 3600
Qy 3773 ATGTCCCTGAATGAATAATCAAGGTGACAGAGCTGTCTCTCTGAAAGGCTCAGAG 3832
Db 3601 ATGTCCCTGAATGAATAATCAAGGTGACAGAGCTGTCTCTCTGAAAGGCTCAGAG 3660
Qy 3833 AGAAGCGGTTCTTGGCATTTTCAGCTTCTAGAGGCTGCTGATTTCCAGGCTCTCAGT 3892
Db 3661 AGAAGCGGTTCTTGGCATTTTCAGCTTCTAGAGGCTGCTGATTTCCAGGCTCTCAGT 3720
Qy 3893 GGTCTGTCAGGCTTTTCTCAATGATGATCATCTGTGACATGCGCCCTCCCATCTCTCT 3952
Db 3721 GGTCTGTCAGGCTTTTCTCAATGATGATCATCTGTGACATGCGCCCTCCCATCTCTCT 3780
Qy 3953 TGAATTTAAAGCCCAAGAGATTCAGAGATTAATCTCTCAATCTTAAGATCTTCTCATC 4012
Db 3781 TGAATTTAAAGCCCAAGAGATTCAGAGATTAATCTCTCAATCTTAAGATCTTCTCATC 3840
Qy 4013 ATCTGGAAGAGCTTTTGCATGCAAGCAACATTAAGCAAGGTGGGATTAAGACAG 4072
Db 3841 ATCTGGAAGAGCTTTTGCATGCAAGCAACATTAAGCAAGGTGGGATTAAGACAG 3900
Qy 4073 GACATCTTTGGGGTCTGTTATTTCTGCTTACCAACACTTCTGCGACTGACTCCAGAG 4132
Db 3901 AACATCTTTGGGGTCTGTTATTTCTGCTTACCAACACTTCTGCGACTGACTCCAGAG 3960
Qy 4133 AGAGGCTACAAAATGATCTGGCGCAAGAGATGTTTTTTAGCTTGGGACTTAAAC 4192
Db 3961 AGAGGCTACAAAATGATCTGGCGCAAGAGATGTTTTTTAGCTTGGGACTTAAAC 4020
Qy 4193 TT-AAAAAACCAGATCAGAAATCTGGCCATGCTGGGGCTCACTTCTCACTAGCA 4251
Db 4021 TTTAAAAAACCAGATCAGAAATCTGGCCATGCTGGGGCTCACTTCTCACTAGCA 4080
Qy 4251 ACACTGGCTGAGGCTGGGCAACAGCTGCTTTAAGAGGGGTCACCTTCAACAGGT 4311
Db 4081 ACACTGGCTGAGGCTGGGCAACAGCTGCTTTAAGAGGGGTCACCTTCAACAGGT 4140
Qy 4312 CACCAAGCCCACTATGAGCCCTATCACTTCCCAATGAGGCTAAGTGTGTTCTAC 4371
Db 4141 CACCAAGCCCACTATGAGCCCTATCACTTCCCAATGAGGCTAAGTGTGTTCTAC 4200
Qy 4372 TGAATGATGCTGTCAGGTCATTTATGTAATGAAAAAGAAAGACTGGATTAATCT 4431
Db 4201 TGAATGATGCTGTCAGGTCATTTATGTAATGAAAAAGAAAGACTGGATTAATCT 4260
Qy 4432 CTAACTAGGTGATGACATGAGCAATGTCATCAATTAATCTCTCTCTCTCTCTCT 4491
Db 4261 CTAACTAGGTGATGACATGAGCAATGTCATCAATTAATCTCTCTCTCTCTCTCTCT 4320
Qy 4492 TCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 4551
Db 4321 TCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 4380
Qy 4552 GAGTCAGTCGCGCAATCTGCGCTCACTGCAACTCTGCTCTGCGGCTCAAGCAATTC 4611
Db 4381 GAGTCAGTCGCGCAATCTGCGCTCACTGCAACTCTGCTCTGCGGCTCAAGCAATTC 4440
Qy 4612 TCCCACTGAGCTCCCAATGTCGAGGATCACTGGCAACCAACCACTGAGCTAT 4671
Db 4441 TCCCACTGAGCTCCCAATGTCGAGGATCACTGGCAACCAACCACTGAGCTAT 4500
Qy 4672 TTTGATTTTGTGAGAGACAGGGTTTCAACATGTTGCGCAGGCTGCTCAACCTCTG 4731
Db 4501 TTTGATTTTGTGAGAGACAGGGTTTCAACATGTTGCGCAGGCTGCTCAACCTCTG 4560
Qy 4732 GGTCAAGCAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4791
Db 4561 GGTCAAGCAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4620
Qy 4792 CGCATTCAGCCCAACCTCTATTTATACCAATTAACCTGCGCAGTAACTGTCGCTTTG 4851
Db 4621 CGCATTCAGCCCAACCTCTATTTATACCAATTAACCTGCGCAGTAACTGTCGCTTTG 4680
Qy 4852 CTTCCTCAACCTGCTGCTGATCTGGAAGAGAGGATTAATGTTATAGCTTGCACAG 4911
Db 4681 CTTCCTCAACCTGCTGCTGATCTGGAAGAGAGGATTAATGTTATAGCTTGCACAG 4740
Qy 4912 TCCCAAGTTCAATTTCTGCGCAAAAATCTTCTTCAAAAATTAATGATCTTCAATGT 4971
Db 4741 TCCCAAGTTCAATTTCTGCGCAAAAATCTTCTTCAAAAATTAATGATCTTCAATGT 4800
Qy 4972 ATTCAATGAATTCACCTTGGAAATGCAACCGCTCAACTTGTTCACATGGAATTAAGAA 5031
Db 4801 ATTCAATGAATTCACCTTGGAAATGCAACCGCTCAACTTGTTCACATGGAATTAAGAA 4860
Qy 5032 GGAATTTATAGTCTCTTAAATGCGGTGATCTGCAAGACCTTGTGAACATTTCCAGAG 5091
Db 4861 GGAATTTATAGTCTCTTAAATGCGGTGATCTGCAAGACCTTGTGAACATTTCCAGAG 4920
Qy 5092 ATAGGATATTTAAGTCATGCTTGGCGTGTGCTAATGACATTTCTCTTGTGAAGTCT 5151
Db 4921 ATAGGATATTTAAGTCATGCTTGGCGTGTGCTAATGACATTTCTCTTGTGAAGTCT 4977
```

[illegible]

DEFINITION	Primer for synthesizing full-length cDNA and use thereof.
ACCESSION	BD160235
VERSION	BD160235.1 GI:27865993
KEYWORDS	JP 2002191363-A/15078.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	Oca,T., Isogai,T., Nishikawa,T., Hayaashi,K., Saito,K., Yamamoto,J., Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
TITLE	Isoli for synthesizing full-length cDNA and use thereof
JOURNAL	Patent: JP 2002191363-A-15078 09-JUL-2002;
COMMENT	HELIX RESEARCH INSTITUTE OS Homo sapiens (human) PN JP 2002191363-A/15078 PD 09-JUL-2002 PF 28-JUL-2000 JP 2000280990 PI TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU SAIRO, PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU, PI KEIICHI NAGAI,TETSUO OTSUKI PC C12N5/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/PC 10, PC C12P21/02,C12P1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC Primer for synthesizing full-length cDNA and use thereof FH Key Location/Qualifiers FT CDS (275)..(2221).
FEATURES	Location/Qualifiers source 1..5573 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" BASE COUNT 1373 a 1524 c 1290 g 1386 t ORIGIN
Query Match	..77.7%; Score 4786; DB 6; Length 5573;
Best Local Similarity	99.8%; Pred. No. 0;
Matches 5566;	Conservative 0; Mismatches 6; Indels 4; Gaps 4;
569	ATATTGACTGATGTCAAGGCCCATGTCTTTTGGGAGGTTCAAGCTTCAGCCCTCGACAT 628
1	ATATTGACTGATGTCAAGGCCCATGTCTTTTGGGAGGTTCAAGCTTCAGCCCTCGACAT 60
629	CGTTGGTGCCTACCCCTCAAGCAAGAACTTCATCTGGGAGTGCAGAAAGCTCATTAAGAGCATCG 688
61	CGTTGGTGCCTACCCCTCAAGCAAGAACTTCATCTGGGAGTGCAGAAAGCTCATTAAGAGCATCG 120
689	GTTTAGAGCTGCAGATTTCATCCCTCGCTCGAGAGCGATCGATCGCGGGTGAAGAGCTGCC 748
121	G-TTAGAGCTGCAGATTTCATCCCTCGCTCGAGAGCGATCGATCGCGGGTGAAGAGCTGCC 179
749	CAGACGAGTCACTCACTCCATCAAGCGGCGGAATGCATGCCACCGTGTCAAGATCGGAA 808
180	CAGACGAGTCACTCACTCCATCAAGCGGCGGAATGCATGCCACCGTGTCAAGATCGGAA 239
809	CGTTCTCAGCAATGGAGCTGTGCTCCGGAGTCAAGATGCAGAAAGAGTGAATAATGGGCT 868
240	CGTTCTCAGCAATGGAGCTGTGCTCCGGAGTCAAGATGCAGAAAGAGTGAATAATGGGCT 299
869	TACACCTCCCATGGTTCCACCCCAAGAAATGTCTCCGAGCTTCAGCATTCAGAAACCGCTCAT 928
300	TACACCTCCCATGGTTCCACCCCAAGAAATGTCTCCGAGCTTCAGCATTCAGAAACCGCTCAT 359
929	CTATTAACAGCTCTGTGATCATGAGTCTGTGTTTGAAGGTGAAGGCTCAGCAACCTCTGA 988
360	CTATTAACAGCTCTGTGATCATGAGTCTGTGTTTGAAGGTGAAGGCTCAGCAACCTCTGA 419
989	TGTGTGCAACTACCCAGAAAGGCTTCCCTGAGATGAGTGCATATGACGTGGCAGTTTGTGC 1048
420	TGTGTGCAACTACCCAGAAAGGCTTCCCTGAGATGAGTGCATATGACGTGGCAGTTTGTGC 479



QY 1049 TTCTGCAACCTGCGGCGCAAGCTCTCTTCTCAACTTCAACTCTTCCAACTGTGAGA 1108  
DB 480 TTCTGCAACCTGCGGCGCAAGCTCTCTTCTCAACTTCAACTCTTCCAACTGTGAGA 539  
QY 1109 GGAAGAGAGAGGGGTGTAATACTACATCCGGGGCTCCACCAACCCCGAGGTGTA 1168  
DB 540 GGAAGAGAGAGGGGTGTAATACTACATCCGGGGCTCCACCAACCCCGAGGTGTA 599  
QY 1169 AGCTGAGAGACAAGCAGCTGGGAGCATGCGGGGAACTTCAACCTCTCTGCAAGGCT 1228  
DB 600 AGCTGAGAGACAAGCAGCTGGGAGCATGCGGGGAACTTCAACCTCTCTGCAAGGCT 659  
QY 1229 GTGACCAAGATGCCCAAGATCCAGGGATCTCCGCGTCAAGTTCGAATTTTGTGCAAC 1288  
DB 660 GTGACCAAGATGCCCAAGATCCAGGGATCTCCGCGTCAAGTTCGAATTTTGTGCAAC 719  
QY 1289 ATCCCAAAATGAAGCAATTAATAATAGTGTGTAATTTAGTAAATGAGCCAGCATGT 1348  
DB 720 ATCCCAAAATGAAGCAATTAATAATAGTGTGTAATTTAGTAAATGAGCCAGCATGT 779  
QY 1349 CACTCAGCATGAGCGGCGGCGTCAAAACAGAGCGCAAGTTTGTCCCTGGCTGTTTGG 1408  
DB 780 CACTCAGCATGAGCGGCGGCGTCAAAACAGAGCGCAAGTTTGTCCCTGGCTGTTTGG 839  
QY 1409 TGTGTCTAGAATCTCGGACCTGACGTGAACCTCACTCAATCTGCTCCAAACACA 1468  
DB 840 TGTGTCTAGAATCTCGGACCTGACGTGAACCTCACTCAATCTGCTCCAAACACA 899  
QY 1469 AAATCTCTTCTCTTTGTGATGATCTGACACGTCTGTGATGAAATGTGAAAAACATTA 1528  
DB 900 AAATCTCTTCTCTTTGTGATGATCTGACACGTCTGTGATGAAATGTGAAAAACATTA 959  
QY 1529 GGTGCAACAGACACGCGTCTGCGCAAAAGAAATCTACTCACTCCAGGAGGCCAGTGA 1588  
DB 960 GGTGCAACAGACACGCGTCTGCGCAAAAGAAATCTACTCACTCCAGGAGGCCAGTGA 1019  
QY 1589 TCTCTCACTGCTGTGAGAGCTGACATGACTTCTCTGAAAGCTGTGAGGCCAAGACA 1648  
DB 1020 TCTCTCACTGCTGTGAGAGCTGACATGACTTCTCTGAAAGCTGTGAGGCCAAGACA 1079  
QY 1649 GGCTCAGCTGTGCTGTGAGAGGCCAGAGAGCTGACAGCATACACAGAAAGCCCT 1708  
DB 1080 GGCTCAGCTGTGCTGTGAGAGGCCAGAGAGCTGACAGCATACACAGAAAGCCCT 1139  
QY 1709 GGAACACGAGCTTACGCTACCTGCTGAGGAGGAGGACATCCAGGCGAGGCTGATCTG 1768  
DB 1140 GGAACACGAGCTTACGCTACCTGCTGAGGAGGAGGACATCCAGGCGAGGCTGATCTG 1199  
QY 1769 GCTCTCTTCTGCGGAGGCTCTATCAAGCAGATCCAGGTGAAGCAGAACATCTCGGTGA 1828  
DB 1200 GCTCTCTTCTGCGGAGGCTCTATCAAGCAGATCCAGGTGAAGCAGAACATCTCGGTGA 1259  
QY 1829 CCCTTTCGACCTTTTGGCCCCAGCTTTCAAACAGAGGCTTCAGGCGAGGCTGACGCTGT 1888  
DB 1260 CCCTTTCGACCTTTTGGCCCCAGCTTTCGAACAGAGGCTTCAGGCGAGGCTGACGCTGT 1319  
QY 1889 CCTTTATACCTTATTTCAAAAGAGAAAGGCTTTTTCAGGTGACCCCTGACAAAAAGCA 1948  
DB 1320 CCTTTATACCTTATTTCAAAAGAGAAAGGCTTTTTCAGGTGACCCCTGACAAAAAGCA 1379  
QY 1949 AGGTCTACCTGAGAGACCCCAACTGAGGAGCGGGGCTGACATCCCTCACTCTGTGTCT 2008  
DB 1380 AGGTCTACCTGAGAGACCCCAACTGAGGAGCGGGGCTGACATCCCTCACTCTGTGTCT 1439  
QY 2009 GGAACATCAGCGTGGCCAGAGACAGGTGGCTGTGCTGACTTTCTTTTAAAGAGCGAGCG 2068  
DB 1440 GGAACATCAGCGTGGCCAGAGACAGGTGGCTGTGCTGACTTTCTTTTAAAGAGCGAGCG 1499  
QY 2069 GCGTGTCTGCGAGAGAGGCGCGCATTTATGATCATTCAGAGACGCGGAGCCCGGGCTG 2128  
DB 1500 GCGTGTCTGCGAGAGAGGCGCGCATTTATGATCATTCAGAGAGCGAGGAGCCCGGGCTG 1559  
QY 2129 AGGAGATCTTTCAGCGCTGAGAGAGATGTCTCCCAAGCGAAGCTTCAACATTCAGAGCT 2188

DB 1560 AGGAGATCTTTCAGCGCTGAGAGAGATGTCTCCCAAGCGAAGCTTCAACATTCAGAGCT 1619  
QY 2189 TCTGGGTCAACATCTCTTAATGACAGCCCAAGAGCGGCAAGAGCTAGACCTGCTTCT 2248  
DB 1620 TCTGGGTCAACATCTCTTAATGACAGCCCAAGAGCGGCAAGAGCTAGACCTGCTTCT 1679  
QY 2249 CGGTGACATTAACCCCAAGAGCTGTGACCTTGATGATCTCTCAATCCGAGCGGTGGAG 2308  
DB 1680 CGGTGACATTAACCCCAAGAGCTGTGACCTTGATGATCTCTCAATCCGAGCGGTGGAG 1739  
QY 2309 GTGAGTCTTACTGCTGTCTGCTCCCTCGGGCTATCATTTGCTGTGTGAAAAAGAAAA 2368  
DB 1740 GTGAGTCTTACTGCTGTCTGCTCCCTCGGGCTATCATTTGCTGTGTGAAAAAGAAAA 1799  
QY 2369 AGAAGACAAAGAGGCGCGGCTGTGGGATATCAAGTGGCAACATCACTAGATGTC 2428  
DB 1800 AGAAGACAAAGAGGCGCGGCTGTGGGATATCAAGTGGCAACATCACTAGATGTC 1859  
QY 2429 CGAGGACGCAAAAAAGTTTCAGAAAGGCGCAAGACATGACTCCCATGTGTATGAC 2488  
DB 1860 CGAGGACGCAAAAAAGTTTCAGAAAGGCGCAAGACATGACTCCCATGTGTATGAC 1919  
QY 2489 TCATGAGAGACCAATGATTATGAGGCAATCTGCTACAGATTCAGCGGCTCTTCTGTC 2548  
DB 1920 TCATGAGAGACCAATGATTATGAGGCAATCTGCTACAGATTCAGCGGCTCTTCTGTC 1979  
QY 2549 AGCCAGAGTGAACACCTAACCGGCGCTTCAGGGACCAATGGGGGTCTGTCTCCCTCC 2608  
DB 1980 AGCCAGAGTGAACACCTAACCGGCGCTTCAGGGACCAATGGGGGTCTGTCTCCCTCC 2039  
QY 2609 CACCCACATATGCTTCAGAGGCGCCCACTGCAAAAGTGGGCACTGAGGAGCCACTCTC 2668  
DB 2040 CACCCACATATGCTTCAGAGGCGCCCACTGCAAAAGTGGGCACTGAGGAGCCACTCTC 2099  
QY 2669 GCTCCCTCTCTGAGTCTGAGATGAACCGTACACCTTCTCCATCCAAATGGGGATG 2728  
DB 2100 GCTCCCTCTCTGAGTCTGAGATGAACCGTACACCTTCTCCATCCAAATGGGGATG 2159  
QY 2729 TAAAGACAGAGACACACATCTCCCTTAATGAACTCAGAGGCCATGAGGCCAGCAG 2788  
DB 2160 TAAAGACAGAGACACACATCTCCCTTAATGAACTCAGAGGCCATGAGGCCAGCAG 2219  
QY 2789 AATACTGATCATTTCCAGAGGCTTGTGATGTTTCAATTAAGAGGGGACATGAGACCC 2848  
DB 2220 AATACTGATCATTTCCAGAGGCTTGTGATGTTTCAATTAAGAGGGGACATGAGACCC 2279  
QY 2849 CGTCCGTGTTCTTAACAGAAATCTTAAGAGAGAAATTAATACAGAGAACAGCAGAG 2908  
DB 2280 CGTCCGTGTTCTTAACAGAAATCTTAAGAGAGAAATTAATACAGAGAACAGCAGAG 2339  
QY 2909 GGTTTTCTGAGACCGCCCACTTCAATGCTCTAGTGAAGCTCAATTTTAAGGGCAAGACA 2968  
DB 2340 GGTTTTCTGAGACCGCCCACTTCAATGCTCTAGTGAAGCTCAATTTTAAGGGCAAGACA 2399  
QY 2969 TTGAATAATGAATTAATTCGAATCTGATACAGTCAAGACATCTGCTCTCACTTA 3028  
DB 2400 TTGAATAATGAATTAATTCGAATCTGATACAGTCAAGACATCTGCTCTCACTTA 2459  
QY 3029 GGTCTGCGTTTAAAGCAGGCTGTAATGAGAGAGAGGCGCTGAGTCACTGAGCTAAGG 3088  
DB 2460 GGTCTGCGTTTAAAGCAGGCTGTAATGAGAGAGAGGCGCGAGTCACTGAGCTAAGG 2519  
QY 3089 TTGACAGAAAGCTGTGATTCAGAGTGTAAACAGAGGCTTGGCTTTCAGGACAAAGT 3148  
DB 2520 TTGACAGAAAGCTGTGATTCAGAGTGTAAACAGAGGCTTGGCTTTCAGGACAAAGT 2579  
QY 3149 TTCAATTCAGAGAGACCTTACAGAGTCCCTTCTCACTGGGGTCCCAAGATGAAGAAC 3208  
DB 2580 TTCAATTCAGAGAGACCTTACAGAGTCCCTTCTCACTGGGGTCCCAAGATGAAGAAC 2639  
QY 3209 GACATGTGCTTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 3268

Dh 2640 GACATGTGCTTTTATTATTATTATTGTTGTCCTGTTATTATTAAGATCAAAAT 2699  
Qy 3269 GATATAACACCTAGCTCTTTTACCTGACTAGTAATACCTACTAACTGGTTGGAT 3328  
Dh 2700 GATATAACACCTAGCTCTTTTACCTGACTAGTAATACCTACTAACTGGTTGGAT 2759  
Qy 3329 GCTGGGTTGACTCTTCTGACTGACCGGTGATTAAGTGTGCGCTGCTCCGAGGTGTG 3388  
Dh 2760 GCTGGGTTGACTCTTCTGACTGACCGGTGATTAAGTGTGCGCTGCTCCGAGGTGTG 2819  
Qy 3389 GAATAATTAACAATCTGTCCCAACAGAAAAGATGTGTGTTTGAGACATTTGACACA 3448  
Dh 2820 GAATAATTAACAATCTGTCCCAACAGAAAAGATGTGTGTTTGAGACATTTGACACA 2879  
Qy 3449 TATCTCTTGTATGAAGACTTCTGATTTCTAGTGTGTTGCTGTTATCCATTTGTG 3508  
Dh 2880 TATCTCTTGTATGAAGACTTCTGATTTCTAGTGTGTTGCTGTTATCCATTTGTG 2939  
Qy 3509 GAAATTCATCTTGAATTCCTTGTCTATGTCTTATGCAATTAAGAAAATTTCTCAAGT 3568  
Dh 2940 GAAATTCATCTTGAATTCCTTGTCTATGTCTTATGCAATTAAGAAAATTTCTCAAGT 2999  
Qy 3569 TTCCATGTGGGTTCTCTGACTGACGAATACTTTGACATTTAAAGAAAATTTGAGA 3628  
Dh 3000 TTCCATGTGGGTTCTCTGACTGACGAATACTTTGACATTTAAAGAAAATTTGAGA 3059  
Qy 3629 AATATCTCATCTCTTAAAAATGTTTAAATATATATACAAAACAGTGGCCCCCTGCATTAGTT 3688  
Dh 3060 AATATCTCATCTCTTAAAAATGTTTAAATATATATACAAAACAGTGGCCCCCTGCATTAGTT 3119  
Qy 3689 TTCTGTGGCACTGCAACCCATTAATTGTGATGTTAAAAACAACATTAAGCTTATAGTC 3748  
Dh 3120 TTCTGTGGCACTGCAACCCATTAATTGTGATGTTAAAAACAACATTAAGCTTATAGTC 3179  
Qy 3749 CTGGGATCAGAAATTCAAAATGATATGTCCTGATATGAAAATCAAGTGTGACAGAGCT 3808  
Dh 3180 CTGGGATCAGAAATTCAAAATGATATGTCCTGATATGAAAATCAAGTGTGACAGAGCT 3239  
Qy 3809 GTGCTCTCTTGTGAAGCTCTAGGAGAAAGCCGGTTCTTGGCAATTTCAAGCTTCTAGAG 3868  
Dh 3240 GTGCTCTCTTGTGAAGCTCTAGGAGAAAGCCGGTTCTTGGCAATTTCAAGCTTCTAGAG 3299  
Qy 3869 CTGGCTGCAATTTCCAGAGCTTCAAGTGTGTCAAGCTTTTCTCAATGTGATACGTGA 3928  
Dh 3300 CTGGCTGCAATTTCCAGAGCTTCAAGTGTGTCAAGCTTTTCTCAATGTGATACGTGA 3359  
Qy 3929 CACTGGCCCTCCCACTTCCCTCTTTGACTTACAAAGGCCACAGAAAGATCCAGATTAAT 3988  
Dh 3360 CACTGGCCCTCCCACTTCCCTCTTTGACTTACAAAGGCCACAGAAAGATCCAGATTAAT 3419  
Qy 3989 CTCTCCATCTAAAGATCTTTCATCTCTGGAAGAGCTTTTGCATGCAAGCAACATA 4048  
Dh 3420 CTCTCCATCTAAAGATCTTTCATCTCTGGAAGAGCTTTTGCATGCAAGCAACATA 3479  
Qy 4049 GCCACAGGTGGGATTAAGACCAAGACATCTTTGGGGTGTGTTATTTGCTTACACAC 4108  
Dh 3480 GCCACAGGTGGGATTAAGACCAAGACATCTTTGGGGTGTGTTATTTGCTTACACAC 3539  
Qy 4109 CTTCCAGCACTGACTCCCAAGAGAGGCTAACAAATGATCTGGGCAACAGGGATGTTT 4168  
Dh 3540 CTTCCAGCACTGACTCCCAAGAGAGGCTAACAAATGATCTGGGCAACAGGGATGTTT 3599  
Qy 4169 TGTTTAGCTTGCAGACTTAACACTT-AAAAAAACCCAGATCAGAAAGATCTGGCCATGC 4227  
Dh 3600 TGTTTAGCTTGCAGACTTAACACTTAAAAAAACCCAGATCAGAAAGATCTGGCCATGC 3659  
Qy 4228 TGGGGCTCACATTTCTACCTAGCAACAATCTGGGTGAGCTGGGCAACAGCTCTGCTTTA 4287  
Dh 3660 TGGGGCTCACATTTCTACCTAGCAACAATCTGGGTGAGCTGGGCAACAGCTCTGCTTTA 3719  
Qy 4288 GAAAGGGGTGCTCACTTCAACAGGTCAACAAGGCCACATGAGCCCTATCATTTCCACA 4347  
Dh 3720 GAAAGGGGTGCTCACTTCAACAGGTCAACAAGGCCACATGAGCCCTATCATTTCCACA 3779

Qy 4348 ATGAGCTTAACTGTTGTTTCTACTGATCATATGCCCTGACAGTTGCATTTATGTAATG 4407  
Dh 3780 ATGAGCTTAACTGTTGTTTCTACTGATCATATGCCCTGACAGTTGCATTTATGTAATG 3839  
Qy 4408 AAAAAAGAACTGGGATTAATCTCTAATCAGGTAGTAGACATGACCAATGATGTC 4467  
Dh 3840 AAAAAAGAACTGGGATTAATCTCTAATCAGGTAGTAGACATGACCAATGATGTC 3899  
Qy 4468 TCACATTAACCTTTTCTTTTTTTTCTTTTCTTTTCTTTTTTTTTTAAATGTAGACA 4527  
Dh 3900 TCACATTAACCTTTTCTTTTTTTTCTTTTCTTTTCTTTTTTTTTTAAATGTAGACA 3959  
Qy 4528 GGAATCTCATTTCTGTTGCTTGAAGCTGAGTGCAGTGGCCGAATCTGGCTCATGCAACT 4587  
Dh 3960 GGAATCTCATTTCTGTTGCTTGAAGCTGAGTGCAGTGGCCGAATCTGGCTCATGCAACT 4019  
Qy 4588 CTGGCCCTGGGCTCAAGCAATTCCTCCACCTGACGCTCCCAATATAGCTGGGATCACTGG 4647  
Dh 4020 CTGGCCCTGGGCTCAAGCAATTCCTCCACCTGACGCTCCCAATATAGCTGGGATCACTGG 4079  
Qy 4648 CACAACCAACATGCCCAGCTAATTTGTATTTTGTAGACAGGGTTTCAACATGTT 4707  
Dh 4080 CACAACCAACATGCCCAGCTAATTTGTATTTTGTAGACAGGGTTTCAACATGTT 4138  
Qy 4708 GCCCAGGCTGTCTCAACCTCTGGGCTCAAGCAATTCCTCTGCTCGGCTCCCAAGT 4767  
Dh 4139 GCCCAGGCTGTCTCAACCTCTGGGCTCAAGCAATTCCTCTGCTCGGCTCCCAAGT 4198  
Qy 4768 GCTGGGATTAAGAGTAGTAGGACCGACCGCAAGCCCAACCTCATTTATACCAATTAG 4827  
Dh 4199 GCTGGGATTAAGAGTAGTAGGACCGACCGCAAGCCCAACCTCATTTATACCAATTAG 4258  
Qy 4828 CTGCCAGTAACTGTGACTTTTGTCTCTCAACCCCTGCTGATCTGGAAGAGAGGA 4887  
Dh 4259 CTGCCAGTAACTGTGACTTTTGTCTCTCAACCCCTGCTGATCTGGAAGAGAGGA 4318  
Qy 4888 TTATGTTATAGCTTGTCAAGACAGTCCCAAGTTCAATATTTCTGGGCAAAATCTTCTT 4947  
Dh 4319 TTATGTTATAGCTTGTCAAGACAGTCCCAAGTTCAATATTTCTGGGCAAAATCTTCTT 4378  
Qy 4948 CAAAAAATAAATGTAATCTTATGTAATGAATTAATCACTTGAATTAAGCAACCGCTCA 5007  
Dh 4379 CAAAAAATAAATGTAATCTTATGTAATGAATTAATCACTTGAATTAAGCAACCGCTCA 4438  
Qy 5008 CTTGTTCAATAGGCAATTAATGAAGAAATTTATAGTCTCTAATAGGCTGTAATGCA 5067  
Dh 4439 CTTGTTCAATAGGCAATTAATGAAGAAATTTATAGTCTCTAATAGGCTGTAATGCA 4498  
Qy 5068 GACTCTTGAACATTTTCCAGAGATAGATATTTAAGTATGCTTGGCGTTGCCCTAT 5127  
Dh 4499 GACTCTTGAACATTTTCCAGAGATAGATATTTAAGTATGCTTGGCGTTGCCCTAT 4558  
Qy 5128 GGCACCTTCCCTTGTGAAGCTGTGCTTCCGCAAGGACCTTGGCTGTGAGCGGA 5187  
Dh 4559 GGCACCTTCCCTTGTGAAGCTGTGCTTCCGCAAGGACCTTGGCTGTGAGCGGA 4618  
Qy 5188 GATGCTGACCTGCAATTAAGGGCCAAAGAGGGCTGCGGCTTCTTCCCTCACTGAAG 5247  
Dh 4619 GATGCTGACCTGCAATTAAGGGCCAAAGAGGGCTGCGGCTTCTTCCCTCACTGAAG 4678  
Qy 5248 CCCTTATTTGAATTGACTGTGTGAGCCCTAGCCCTGCAATTTCTGACATTTCCCAACCTC 5307  
Dh 4679 CCCTTATTTGAATTGACTGTGTGAGCCCTAGCCCTGCAATTTCTGACATTTCCCAACCTC 4738  
Qy 5308 CCAAGCCCTTCCAGACAGAGTAAGTGTGCTGCAATTTCCACCAAGTGGATTTGGCTTC 5367  
Dh 4739 CCAAGCCCTTCCAGACAGAGTAAGTGTGCTGCAATTTCCACCAAGTGGATTTGGCTTC 4797  
Qy 5368 CTTAGGCTGGCTACTGTCAACATCAACGACATCACTGTGGCTGCAAGAGACCAACGCTG 5427  
Dh 4798 CTTAGGCTGGCTACTGTCAACATCAACGACATCACTGTGGCTGCAAGAGACCAACGCTG 4857



Qy	5428	GCACATTTTCCCTTCAACTAGAGGAGCTCAAAACATCTCGAGCAAGTGGTGGCTCTAGACCA	5487
Db	4858	GCCATTTTCCCTTCAACTAGAGGAGCTCAAAACATCTCGAGCAAGTGGTGGCTCTAGACCA	4917
Qy	5488	GTATTTTCCGTGAGCTGTGCTCAGTGAAGGGGGCCAGCCTGAGAACCCCTGGCTCTTTTC	5547
Db	4918	GTATTTTCCGTGAGCTGTGCTCAGTGAAGGGGGCCAGCCTGAGAACCCCTGGCTCTTTTC	4977
Qy	5548	TTTAAAGCCCAAGGCCCTTACATATAAACTTTTCAAGGGTCACTGGAACAGTGAAGTGC	5607
Db	4978	TTTAAAGCCCAAGGCCCTTACATATAAACTTTTCAAGGGTCACTGGAACAGTGAAGTGC	5037
Qy	5608	CATTGTGGAAGCCTACTGATGACGAGCCCACTGCTATTCACAGTGGCTGCGCATGGCTA	5667
Db	5038	CATTGTGGAAGCCTACTGATGACGAGCCCACTGCTATTCACAGTGGCTGCGCATGGCTA	5097
Qy	5668	CGAAGAGGCCAGGCGCATGACAGACTGTGCTCTATATGCTGTGATTCATGACAGAGAGGA	5727
Db	5098	CGAAGAGGCCAGGCGCATGACAGACTGTGCTCTATATGCTGTGATTCATGACAGAGAGGA	5157
Qy	5728	AAGGCTTCAGAGAAAGATCAACTGGGACAGACACAGCCACCGGACATGGCCTTGGTAA	5787
Db	5158	AAGGCTTCAGAGAAAGATCAACTGGGACAGACACAGCCACCGGACATGGCCTTGGTAA	5217
Qy	5788	AGGTTAGCAGACTGGTGTGTGTGATCTGACAGTGGCTTACATGGAATTAATTATCATTTG	5847
Db	5218	AGGTTAGCAGACTGGTGTGTGTGATCTGACAGTGGCTTACATGGAATTAATTATCATTTG	5277
Qy	5848	CAGATACTCTTTTAAAGTGGCATTTTATTCATTTCCGTGCTTTAAATAACAATATGACCA	5907
Db	5278	CAGATACTCTTTTAAAGTGGCATTTTATTCATTTCCGTGCTTTAAATAACAATATGACCA	5337
Qy	5908	AAAAACAAGTATCAAGCTGTTTAAAGTGGCTTACCTTGTCCCTGGTTCAGTAGAGGC	5967
Db	5338	AAAAACAAGTATCAAGCTGTTTAAAGTGGCTTACCTTGTCCCTGGTTCAGTAGAGGC	5397
Qy	5968	CCCGGTTTCCCAAGTGTGACTGTGACAGGCTCAGACATGGGCTCAGCAGATGCTGTCTTA	6027
Db	5398	CCCGGTTTCCCAAGTGTGACTGTGACAGGCTCAGACATGGGCTCAGCAGATGCTGTCTTA	5457
Qy	6028	ATTGTGGATGATACAGAAAGCCAGGCTTGGGATACAAAGTGTCTTCCCTTCATTTGAT	6087
Db	5458	ATTGTGGATGATACAGAAAGCCAGGCTTGGGATACAAAGTGTCTTCCCTTCATTTGAT	5517
Qy	6088	GCCGTGCACTGTGTGAAGCAGATGTTTTGTCCGGAAATAAATAATATAGTCTTGG	6143
Db	5518	GCCGTGCACTGTGTGAAGCAGATGTTTTGTCCGGAAATAAATAATATAGTCTTGG	5573
RESULT 6			
AK023834		5573 bp mRNA linear PRI 01-AUG-2002	
LOCUS	AK023834		
DEFINITION	Homo sapiens cDNA FLJ13772 f.18, clone PLACE000300.		
ACCESSION	AK023834		
VERSION	AK023834.1 GI:10435889		
KEYWORDS	oligo capping; f18 (full insect sequence).		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE			
AUTHORS	1 Isegai,T., Oca,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y., Mishikawa,T., Nagai,K., Sugano,S., Takahashi-Fujii,A., Hara,H., Tanase,T., Nomura,Y., Togiya,S., Komai,F., Hara,R., Takeuchi,K., Arita,M., Nabekura,T., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagahari,K., Masuno,Y. and Oshima,A. NEBO human cDNA sequencing project Unpublished 2 (bases 1 to 5573) Isegai,T. and Otsuki,T. Direct Submission Submitted (23-AUG-2000) Takao Isegai, Helix Research Institute, Genomics Laboratory, 1532-3Yana, Kisarazu, Chiba 292-0812, Japan		

**COMMENT**

(E-mail: [gemetics@rii.co.jp](mailto:gemetics@rii.co.jp), Tel: 81-438-52-3975, Fax: 81-438-52-3986).  
NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5'- and 3'-end one pass sequencing and clone selection: Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.

FEATURES  
SOURCE

**CDS**

```
/note="unnamed protein product"  
/codon_start=1
```

BASE COUNT	1373	a	1524	c	1290	g	1386	t
ORIGIN								

Query Match	77.7%;	Score 4786;	DB 9;	Length 5573;
-------------	--------	-------------	-------	--------------

Best Local Similarity 99.8%; Pred.No. 0;  
Matches 5566; Conservative 0; Mismatches 6; Indels 4; Gaps 4;

569 ATATTGACTGTATGTCAGGCCCATGTCCTTTTGGGAGGTTTCAGCTTCAGCCCTCGACAT 628

1 AATATGACCTGATGTCAGGCCAATGTCCTTTGGGAGGTTCAGCTTCAGCCCTCGAAT 60

023 061 061 030 CACCC 1 CACACACAC 1 1 CAC 1 060 061 061 CACAC 1 CACAC 1 CACACACAC 1 C 000

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

\_\_\_\_\_

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1

180 220

809 CAGTCTGACGCAATTGGCACTGTGTCCCGGATCAAGATGCCAAGAAGGAGTGAATAATGGCCT 868

240 CCTTCTGCAGCAATGGCACTGTGTCCCGATCAAGATGCAAGAGGAGTGAAAATGGCCT 299

869 TACACCTCCCATGGTTCCACCACCAGAATGTCCTCCGGCTTCAGCATTGCAAACCGCTCAT 928

300 TACACCTCCATGTTCCACCCAGAAATGTTCTCCGGCTTCAGCATTGCAACCGCTCAT 359

929 CTATAAACGCTCTGTCATCATCGAGTCTGTGTTTGAGGGTGAAGGCTCAGCAACCTGA 988

360 CTATAAACGTCCTGTCATCATCGAGTCCTGCTTGAGCGTGAAGCCTCAGCCATCCCTGA 419

[illegible]

QY 1049 TTCTGCAACCTGCGGGCCAGCGTCTCTTCTCAACTTCAACTCTTCCAACTGTGAGA 1108  
DB 480 TTCTGCAACCTGCGGGCCAGCGTCTCTTCTCAACTTCAACTCTTCCAACTGTGAGA 539  
QY 1109 GGAAGAGAGAGCGGTTGAATACTATACCCGGGCTCCACCAACCCCGAGGTTCGA 1168  
DB 540 GGAAGAGAGAGCGGTTGAATACTATACCCGGGCTCCACCAACCCCGAGGTTCGA 599  
QY 1169 AGCTGAGAGACAAGCAGCTGGGACATGAGCGGGAACTTCAACTCTCTGTGCAAGGCT 1228  
DB 600 AGCTGAGAGACAAGCAGCTGGGACATGAGCGGGAACTTCAACTCTCTGTGCAAGGCT 659  
QY 1229 GTGACCAAGATGCCCAAGTCCAGGGATCTCCGGCTGCACTTCCAGTTTGTTCAC 1288  
DB 660 GTGACCAAGATGCCCAAGTCCAGGGATCTCCGGCTGCACTTCCAGTTTGTTCAC 719  
QY 1289 ATCCACAAATGAACATTAATAATCTAGTGGTTGATCTTGAATGAGCGGCAATGT 1348  
DB 720 ATCCACAAATGAACATTAATAATCTAGTGGTTGATCTTGAATGAGCGGCAATGT 779  
QY 1349 CACTCAACATCGAGCCAGCCCGCTCAAAAGAGCCGCAAGTTTGTCCCTGGCTGTTTGC 1408  
DB 780 CACTCAACATCGAGCCAGCCCGCTCAAAAGAGCCGCAAGTTTGTCCCTGGCTGTTTGC 839  
QY 1409 TGTGTCTAAGATCTCGGACTTGCAATGACAACTTCACTGATCTGGCTCCAAACAA 1468  
DB 840 TGTGTCTAAGATCTCGGACTTGCAATGACAACTTCACTGATCTGGCTCCAAACAA 899  
QY 1469 AATCTCTCTTCTTGTGATGATCTGACAGTGTGAGATGAAATGAGAAAAACATAA 1528  
DB 900 AATCTCTCTTCTTGTGATGATCTGACAGTGTGAGATGAAATGAGAAAAACATAA 959  
QY 1529 GGTGCAACAGACACCGGTACTGCAAAAGAAATCTTACTCACTCCAGGTGCCAGTGA 1588  
DB 960 GGTGCAACAGACACCGGTACTGCAAAAGAAATCTTACTCACTCCAGGTGCCAGTGA 1019  
QY 1589 TCTTCAACCTGCTGTGAGAGTGCATGACTTCTCTGGAAGCTGTGTGCCAAGACA 1648  
DB 1020 TCTTCAACCTGCTGTGAGAGTGCATGACTTCTCTGGAAGCTGTGTGCCAAGACA 1079  
QY 1649 GGTCAAGCTGTGAGTGCAGGCCAGAGCTGAGAGCAATACAGAGAAAGCCCT 1708  
DB 1080 GGTCAAGCTGTGAGTGCAGGCCAGAGCTGAGAGCAATACAGAGAAAGCCCT 1139  
QY 1709 GCAACACAGCTTCACTACCTTGTGAGCAAGTGCATACCCAGCAGGACCTGTACTTGC 1768  
DB 1140 GCAACACAGCTTCACTACCTTGTGAGCAAGTGCATACCCAGCAGGACCTGTACTTGC 1199  
QY 1769 GCTCTCTTGTGCGGAGAGGCTTATCAAGCAATCCAGGTGAAGCAAGATCTCGGTGA 1828  
DB 1200 GCTCTCTTGTGCGGAGAGGCTTATCAAGCAATCCAGGTGAAGCAAGATCTCGGTGA 1259  
QY 1829 CCGTTGGACCTTGTGCGCCAGTTCCAAAGAGGCTCCAGAGGGGTGAGCGGTGT 1888  
DB 1260 CCGTTGGACCTTGTGCGCCAGTTCCAAAGAGGCTCCAGAGGGGTGAGCGGTGT 1319  
QY 1889 CCTTTATACCTTATTTCAAGAGAGAGGCTTTTCAAGGTGACCCCTGACAAAAAGCA 1948  
DB 1320 CCTTTATACCTTATTTCAAGAGAGAGGCTTTTCAAGGTGACCCCTGACAAAAAGCA 1379  
QY 1949 AGGTCTACCTGAGAGACCCCACTGGAGCCGGGGCTGCGCATCTGTGTCTCT 2008  
DB 1380 AGGTCTACCTGAGAGACCCCACTGGAGCCGGGGCTGCGCATCTGTGTCTCT 1439  
QY 2009 GGAACATTCAGCGGCCAGAGACAGGTGCGCTGCACTTTCTTTAAGAGACGAGAG 2068  
DB 1440 GGAACATTCAGCGGCCAGAGACAGGTGCGCTGCACTTTCTTTAAGAGACGAGAG 1499  
QY 2069 GCGTGTCTGCGCAGACAGGGCGGCACTTCAATGATCTCAGAGACAGCGGAGCTG 2128  
DB 1500 GCGTGTCTGCGCAGACAGGGCGGCACTTCAATGATCTCAGAGACAGCGGAGCTG 1559  
QY 2129 AGGAGATCTTCAAGCTGAGCAGAGATGTGCTCCCAAGCCAAAGCTTCCACATCAGCT 2188

DB 1560 AGGAGATCTTCAAGCTGAGCAGAGATGTGCTCCCAAGCCAAAGCTTCCACATCAGCT 1619  
QY 2189 TCTGGGTCAACATCTCTAATGACAGCCCAAGAGCGGCAAGAGCTAGACTGTCTTCT 2248  
DB 1620 TCTGGGTCAACATCTCTAATGACAGCCCAAGAGCGGCAAGAGCTAGACTGTCTTCT 1679  
QY 2249 CGGTGACATTAACCCCAAGAGCTGTGACCTGTATCTCTCATGCGAGGTGGAG 2308  
DB 1680 CGGTGACATTAACCCCAAGAGCTGTGACCTGTATCTCTCATGCGAGGTGGAG 1739  
QY 2309 GTGAGTCTTACTGCTGTGCTCGGCTCATATGATTTGCTGTGAAAAAGAAAAA 2368  
DB 1740 GTGAGTCTTACTGCTGTGCTCGGCTCATATGATTTGCTGTGAAAAAGAAAAA 1799  
QY 2369 AGAAGACAAACAAAGGAGCCCGCTGGGATATCAATGAGCAATCAATACTAGATGC 2428  
DB 1800 AGAAGACAAACAAAGGAGCCCGCTGGGATATCAATGAGCAATCAATACTAGATGC 1859  
QY 2429 CGAGGACGCAAAAAAGTTTCAAAAAAGCGAAGACATGACTTCCATGTGTATGAG 2488  
DB 1860 CGAGGACGCAAAAAAGTTTCAAAAAAGCGAAGACATGACTTCCATGTGTATGAG 1919  
QY 2489 TCATGAGAGACACATGATATGAGGACATGCTACAGGATTCAGCGGCTCTCTG 2548  
DB 1920 TCATGAGAGACACATGATATGAGGACATGCTACAGGATTCAGCGGCTCTCTG 1979  
QY 2549 ACCAGAGGTGACACATACCGGCGTTCCAGGGACACATGAGGGGTCTGCTCTCTCC 2608  
DB 1980 ACCAGAGGTGACACATACCGGCGTTCCAGGGACACATGAGGGGTCTGCTCTCTCC 2039  
QY 2609 CACCAACATATGCTTCAGAGGCGCCCACTGCAAGTTGGCACTGAGAGCCACTCTC 2668  
DB 2040 CACCAACATATGCTTCAGAGGCGCCCACTGCAAGTTGGCACTGAGAGCCACTCTC 2099  
QY 2669 GCTCCCTCTGAGTCTGAGATGAGTAAACGTAACCTTCTCCATCCCAATGGGGATG 2728  
DB 2100 GCTCCCTCTGAGTCTGAGATGAGTAAACGTAACCTTCTCCATCCCAATGGGGATG 2159  
QY 2729 TAAAGACAGAGACACAGACTTCCCTTACTGAACTCAGAGGCCATGAGCCAGCAG 2788  
DB 2160 TAAAGACAGAGACACAGACTTCCCTTACTGAACTCAGAGGCCATGAGCCAGCAG 2219  
QY 2789 AATTAATGATCATTTCCAGAGCGCTTGTGAGTTTCAATAAGAGGGAACCTGAGACCC 2848  
DB 2220 AATTAATGATCATTTCCAGAGCGCTTGTGAGTTTCAATAAGAGGGAACCTGAGACCC 2279  
QY 2849 CGTCCGTGTTCTTAACAGAAATCTTAAGAGAGAAATTAACAGAAAGACAGAGGA 2908  
DB 2280 CGTCCGTGTTCTTAACAGAAATCTTAAGAGAGAAATTAACAGAAAGACAGAGGA 2339  
QY 2909 GGTTCCTGAGACACCGCACTTCACTGCTCAGTGAATCTTAAGGCAAGAGA 2968  
DB 2340 GGTTCCTGAGACACCGCACTTCACTGCTCAGTGAATCTTAAGGCAAGAGA 2399  
QY 2969 TTGAATAATGATTAATTCATCTGATACAGTCAAGCTCATGTGCTCTCAACTTA 3028  
DB 2400 TTGAATAATGATTAATTCATCTGATACAGTCAAGCTCATGTGCTCTCAACTTA 2459  
QY 3029 GGTGTGCGGTTAAGCCAGCTGTATTAAGAGAGAGGCTGAGTCACTGAGCATAGG 3088  
DB 2460 GGTGTGCGGTTAAGCCAGCTGTATTAAGAGAGAGGCTGAGTCACTGAGCATAGG 2519  
QY 3089 TTGACAGCAAGCCCTGATTCAGAGTGTAAACAGAGGCTTGGCTTTGAGAGCAACAGT 3148  
DB 2520 TTGACAGCAAGCCCTGATTCAGAGTGTAAACAGAGGCTTGGCTTTGAGAGCAACAGT 2579  
QY 3149 TCCAAATTCAGAGAGCTTACAGAGTCCCTACTCATGAGGTCCCAAGATGAAGAAC 3208  
DB 2580 TCCAAATTCAGAGAGCTTACAGAGTCCCTACTCATGAGGTCCCAAGATGAAGAAC 2639  
QY 3209 GACATGTGCTTTTATTTATTTATTTTGGTGTCTGTGTATTTAAGAGATCAAAAT 3268

Db	2640	AACAATGCGCTTTTATATATATTTATTTATTTGGTGGTCTGTGTTATTTAAGAAATCAAAAT	2659
Oy	3289	GTATTAACCACTAGCTCTTTTTCACCTGACTTAAGTAATTAATCACTAATCACTAGCTTTGGAT	3328
Db	2700	GTATTAACCACTAGCTCTTTTTCACCTGACTTAAGTAATTAATCACTAATCACTAGCTTTGGAT	2759
Oy	3329	GCCGGGGTGTGACCTTCTTACTGACCGGCTAAGTAACGTTGCTCTGCTCCCAAGGTGGG	3388
Db	2760	GCCGGGGTGTGACCTTCTTACTGACCGGCTAAGTAACGTTGCTCTGCTCCCAAGGTGGG	2819
Oy	3389	GAATTAATTAACAACTGTCTCAACACGAAGAAAGATGTGTGTTTGGAGACGATTTGACACA	3448
Db	2820	GAATTAATTAACAACTGTCTCAACACGAAGAAAGATGTGTGTTTGGAGACGATTTGACACA	2879
Oy	3449	TATCTGCTTTGATAAAGACATTTCTGTATCTCTAGTGGTTCGTGGTTATCCCATTTGTG	3508
Db	2880	TATCTGCTTTGATAAAGACATTTCTGTATCTCTAGTGGTTCGTGGTTATCCCATTTGTG	2939
Oy	3509	GAATTAATCACTTGAATGCCATTTGCTTATCTTACCTTAAGCAATTAAGAAATTTCTCAAGT	3568
Db	2940	GAATTAATCACTTGAATGCCATTTGCTTATCTTACCTTAAGCAATTAAGAAATTTCTCAAGT	2999
Oy	3569	TTCCATGTCGGGTTCCTCTAGCTGACGAATACTTTGACATTTTAAAGAAATTTAGAGA	3628
Db	3000	TTCCATGTCGGGTTCCTCTAGCTGACGAATACTTTGACATTTTAAAGAAATTTAGAGA	3059
Oy	3629	ATATTTCTCATCCTCTTAAATAATGTTTAAATATATATACCAACAGTGGCCCCCTGCATTAAGT	3688
Db	3060	ATATTTCTCATCCTCTTAAATAATGTTTAAATATATATACCAACAGTGGCCCCCTGCATTAAGT	3119
Oy	3689	TTCTGTGGCCACTGCAACCCATTAATCTTGTAGCTTTAAACAACACATTAAGCTTAATAGTC	3748
Db	3120	TTCTGTGGCCACTGCAACCCATTAATCTTGTAGCTTTAAACAACACATTAAGCTTAATAGTC	3179
Oy	3749	CTGGGGATCAGAAATTCGAAATAATGATGTGCTCCGTAATGAATAATCAAGGTGACAGAGACT	3808
Db	3180	CTGGGGATCAGAAATTCGAAATAATGATGTGCTCCGTAATGAATAATCAAGGTGACAGAGACT	3239
Oy	3809	GTGCTCCTTCTGAAAGGCTCTAAGGAGAAAGCCGGTTCCTTGCATTTCAAGCTTCTAGAGG	3868
Db	3240	GTGCTCCTTCTGAAAGGCTCTAAGGAGAAAGCCGGTTCCTTGCATTTCAAGCTTCTAGAGG	3299
Oy	3869	CTGGCTGATTTCCACAGGCTCCAGTGTGGCTGTCAAGCTTTTCTCACATGGCATCACTGTGA	3928
Db	3300	CTGGCTGATTTCCACAGGCTCCAGTGTGGCTGTCAAGCTTTTCTCACATGGCATCACTGTGA	3359
Oy	3929	CACGTGGCCCTCCCACTTGCCTTCTTGAATCTTCAAAAGCCACACAGGAATAATCAGAGTAAT	3988
Db	3360	CACGTGGCCCTCCCACTTGCCTTCTTGAATCTTCAAAAGCCACACAGGAATAATCAGAGTAAT	3419
Oy	3989	CTTCTCATCTTAAAGATCCTTCAATCATCCTGTGAAGAGCCTTTTGGCATGCAAGACAATA	4048
Db	3420	CTTCTCATCTTAAAGATCCTTCAATCATCCTGTGAAGAGCCTTTTGGCATGCAAGACAATA	3479
Oy	4049	GCCACAGGTGGGATTTAGAACCAAGACATCTTTGGGGTGTGTTATTTCTGCTTACCAAC	4108
Db	3480	GCCACAGGTGGGATTTAGAACCAAGACATCTTTGGGGTGTGTTATTTCTGCTTACCAAC	3539
Oy	4109	CTTCTGCGCACTGACTCCCAACAGGAGGCTTACAAATGATCTGGCGGACAGGAGTGTTT	4168
Db	3540	CTTCTGCGCACTGACTCCCAACAGGAGGCTTACAAATGATCTGGCGGACAGGAGTGTTT	3599
Oy	4169	TGTTTAGCTTGCAGACTCTTAAACATT -AAAAAAACCAGATTCAGAAATCTGGCCATGC	4227
Db	3600	TGTTTAGCTTGCAGACTCTTAAACATTAAABAAAAACCAGATTCAGAAATCTGGCCATGC	3659
Oy	4228	TGGGGCTCAATTTCTCACCTAGCAACAATCTGGCTGGAGCTGGGCAACAGCTCTGCTTTA	4287
Db	3660	TGGGGCTCAATTTCTCACCTAGCAACAATCTGGCTGGAGCTGGGCAACAGCTCTGCTTTA	3719
Oy	4288	GAAGGGGTGTCACTTTCACCAAGTGTACCAACAGCCCACTAAGGCCCTTATCTTCCACA	4347
Db	3720	GAAGGGGTGTCACTTTCACCAAGTGTACCAACAGCCCACTAAGGCCCTTATCTTCCACA	3779

[illegible]

```

QY 5428 GCCATTTTCCTTCACTGAGGGCTCAAAAACCTCTGAGCAAGTTGGCTCTCTGAGACCA 5487
DB 4858 GCCATTTTCCTTCACTGAGGGCTCAAAAACCTCTGAGCAAGTTGGCTCTCTGAGACCA 4917
QY 5488 GATATTTCCGAGAGCTGCTCAGTAAGGGGCGCCGCTGAGGAACCTGGCTCTTTTC 5547
DB 4918 GATATTTCCGAGAGCTGCTCAGTAAGGGGCGCCGCTGAGGAACCTGGCTCTTTTC 4977
QY 5548 TTTAAAGCCAGGCGCCCACTTACATATAAAATTTTCAGGGTCACTGGAAAACAGTGAAGTGC 5607
DB 4978 TTTAAAGCCAGGCGCCCACTTACATATAAAATTTTCAGGGTCACTGGAAAACAGTGAAGTGC 5037
QY 5608 CATTTTGAAGCTTACTGATGCCAGCCCACTGCTCATTCACGTGTGCTGCTCCATGCCCTA 5667
DB 5038 CATTTTGAAGCTTACTGATGCCAGCCCACTGCTCATTCACGTGTGCTGCTCCATGCCCTA 5097
QY 5668 CGAGAAAGGCGCAGGATGAGGACTGCTCTAATAGCTGTGGTCACTTGCACAGAAAGGA 5727
DB 5098 CGAGAAAGGCGCAGGATGAGGACTGCTCTAATAGCTGTGGTCACTTGCACAGAAAGGA 5157
QY 5728 AAGGTCTCAAGAAAGAGTCAACTGGGACAAAGCAAGCCCAACCGGACATGCGCTTGTGTA 5787
DB 5158 AAGGTCTCAAGAAAGAGTCAACTGGGACAAAGCAAGCCCAACCGGACATGCGCTTGTGTA 5217
QY 5788 AAGTTCAGACTGTGTGTGTGCTGATCTGCAAGTCTTCACTGGAATAATATTATTCATTG 5847
DB 5218 AAGTTCAGACTGTGTGTGTGCTGATCTGCAAGTCTTCACTGGAATAATATTATTCATTG 5277
QY 5848 CAGATCTTTTAAAGTGGGCAATTTATTCATTTCTGCTGCTTAAATAACAAATGACCA 5907
DB 5278 CAGATCTTTTAAAGTGGGCAATTTATTCATTTCTGCTGCTTAAATAACAAATGACCA 5337
QY 5908 AAAAACAAGTATCAACTGTTTAAAGTCTTGGCTACTTGTCCCTGCTGCTAGTAGAGC 5967
DB 5338 AAAAACAAGTATCAACTGTTTAAAGTCTTGGCTACTTGTCCCTGCTGCTAGTAGAGC 5397
QY 5968 CCGGCTTCCCACTGTTGATCTGTGACAGGCTCAGCATGGGCTCAGCATGCTGTCTTA 6027
DB 5398 CCGGCTTCCCACTGTTGATCTGTGACAGGCTCAGCATGGGCTCAGCATGCTGTCTTA 5457
QY 6028 ATTGTGAGTATGATACGAAAGCCAGGCTTGGATACAGTTCTTCTTCTTCAATTGAT 6087
DB 5458 ATTGTGAGTATGATACGAAAGCCAGGCTTGGATACAGTTCTTCTTCTTCAATTGAT 5517
QY 6088 GCGGCTCACTGTGTGTAAGCAGATGTTTGTCCGGAATAATATATAGTCTTGG 6143
DB 5518 GCGGCTCACTGTGTGTAAGCAGATGTTTGTCCGGAATAATATATAGTCTTGG 5573

```

RESULT 7  
AC105902/c 163115 bp DNA linear PRI 27-JUN-2002  
LOCUS Homo sapiens chromosome 3 clone RP11-578F5, complete sequence.  
DEFINITION AC105902  
ACCESSION AC105902.2 GI:21617714  
VERSION HTG.  
KEYWORDS Homo sapiens (human)  
SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 163115)  
AUTHORS Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,  
Saenphimmachak, C., Phelps, K.A., Buckley, D., Kibukawa, M., Raymond, C.  
and Haugen, E.D.  
TITLE Direct Submission  
JOURNAL Unpublished  
AUTHORS 2 (bases 1 to 163115)  
Kaul, R.K., Olson, M.V., Raymond, C. and Haugen, E.D.  
TITLE Direct Submission  
JOURNAL Submitted (10-JAN-2002) Genome Center, University of Washington,  
Box 352145, Seattle, WA 98195, USA  
REFERENCE 3 (bases 1 to 163115)  
AUTHORS Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,

TITLE  
 JOURNAL  
 COMMENT  
 Saenphimmachak, C., Phelps, K.A., Buckley, D., Kibukawa, M., Raymond, C.  
 and Haugen, E.D.  
 Direct Submission  
 Submitted (27-JUN-2002) Genome Center, University of Washington,  
 Box 352145, Seattle, WA 98195, USA  
 On Jun 27, 2002 this sequence version replaced gi:18104827.

----- Genome Center  
 Center: University of Washington Genome Center  
 Center Code: UMG  
 Web site: <http://www.genome.washington.edu>  
 Contact: [umgchigs@u.washington.edu](mailto:umgchigs@u.washington.edu)  
 ----- Project Information

----- Summary Statistics  
 Center project name: chr-3  
 Center clone name: RP11-578F5 (bc0477)  
 Sequencing vector: plasmid, 100% of reads  
 Chemistry: Dye-terminator ET; 29% of reads  
 Chemistry: Dye-terminator Big Dye; 71% of reads  
 Assembly program: Phrap; version 0.990319  
 Consensus quality: 163063 bases at least Q40  
 Consensus quality: 163110 bases at least Q30  
 Consensus quality: 163115 bases at least Q20  
 Insert size: 163115; sum-of-contrigs  
 Quality coverage: 9.1x in Q20 bases; sum-of-contrigs

----- Overlapping Sequences:  
 5': RP11-68104 (UMGC:bc0528) AC104165  
 3': RP11-487J21 AC010170, 50065-bp overlap

----- Sequence Quality Assessment:  
 This entry has been annotated with sequence quality  
 estimates computed by the Phrap assembly program.  
 All manually edited bases have been reduced to quality zero.  
 Quality levels above 40 are expected to have less than  
 1 error in 10,000 bp.  
 Base-by-base quality values are not generally visible from the  
 Genbank flat file format but are available as part  
 of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted:  
 all regions were either double-stranded or sequenced with an  
 alternate chemistry or covered by high quality data (i.e., Phred  
 quality >= 30); an attempt was made to resolve all sequencing  
 problems, such as compressions and repeats; all regions were  
 covered by at least one plasmid subclone or more than one M13  
 subclone; and the assembly was confirmed by restriction digest.

#### Sequence Validation:

This sequence has been validated by Multiple Complete Digest  
 fingerprinting. Comparison of the experimentally derived digest  
 fragments with sequence-predicted fragments is given below.  
 The electronically-digested sequence consists of both insert and  
 vector, in order to accurately represent the entire circular BAC.  
 Small fragments below a variable cutoff (approximately 400-800 bp)  
 are not resolved in the fingerprint and hence do not appear  
 in the table. There are no significant remaining discrepancies  
 between the experimental and predicted values. Uniquely ordered  
 fragments are separated by dashed lines.

SeqderMap Fingerprint	HindIII	BglII	SeqderMap Fingerprint
8696	8823	3091	3097
6	<800	6382	6416
740	789	512	<800
10200	10022	449	<800
1644	1614	5235	5065
			1052
			1030

9393	9315	2845	2873	852	842
871	897	7321	7256	1202	1151
797	789	5620	5480	18054	18421
15170	15087	1720	1661	2066	2097
919	897	3901	4210	3551	3566
2182	2206	9970	9927	7798	7835
214	<800	3602	3542	1844	1804
6473	6579	522	<800	1354	1284
1456	1431	4155	4210	358	<800
682	<800	3891	3866	291	<800
3675	3703	3267	3284	9862	9710
3439	3449	50	<800	2935	3011
10009	10022	4248	4210	3034	3011
6068	6125	1035	1039	262	<800
11400	11202	3002	2986	6084	6129
96	<800	13216	13421	1769	1804
4100	4108	2978	2986	5507	5488
779	789	8146	8119	144	<800
1221	1181	4465	4507	1459	1448
278	<800	2239	2257	70	<800
8458	8511	45	<800	3690	3725
12014	11848	4368	4507	102	<800
645	<800	286	<800	511	<800
1177	1181	4633	4819	1903	1951
153	<800	1462	1437	5124	5172
1993	1940	2058	2083	367	<800
3153	3167	745	765	7100	7161
132	<800	1321	1310	5890	5897
4939	4944	3528	3542	1814	1804
466	<800	1640	1661	125	<800
1111	1295	8779	8805	43	<800
3183	3167	5089	5065	117	<800
16952	17153	2571	2605	1955	1951
1629	1614	2440	2393	3477	3566
7261	7251	1682	1661	4899	4924
7840	7927	1660	1661	774	760

QY	2409	CAACATCAATCTGAGATGCCAGGCAAGCAAAAGTTTCAGAAAGGCGAAAGACAA	4912	5065	10749	10446
DB	74158	CAACATCAATCTGAGATGCCAGGCAAGCAAAAGTTTCAGAAAGGCGAAAGACAA	2470	2525	1620	1587
QY	2469	TGACTCCCATGTGTATGATGATGAGACACATGTATATGGCATCTGCTACAGGA	8892	8805	8868	8724
DB	74098	TGACTCCCATGTGTATGATGATGAGACACATGTATATGGCATCTGCTACAGGA	1257	1228	5884	5897
QY	2529	TTCCAGCGCTCTTCTCTGACAGGAGTGACACCTACCGGCGTTCCAGGCAACCT	10114	9927	1521	1448
DB	74038	TTCCAGCGCTCTTCTCTGACAGGAGTGACACCTACCGGCGTTCCAGGCAACCT	771	760		
QY	2589	GGGGGTCTGTCTCTCCCTCCGCCACCAATATGCTCCAGGGGCCCACTGCAAGTTGAC	11706	11346		
DB	73978	GGGGGTCTGTCTCTCTCCGCCACCAATATGCTCCAGGGGCCCACTGCAAGTTGAC	369	<800		
QY	2649	CACTGAGAGCCACTCTCTCGCTCCCTCTAGTCTGAGTGAACCGTACACCTTCTC	5177	5172		
DB	73918	CACTGAGAGCCACTCTCTCGCTCCCTCTAGTCTGAGTGAACCGTACACCTTCTC	2916	3011		
QY	2709	CCATCCCAACATGGGATGTATAGACAGACAGACATCTCCCTTAAGTAACTCTA	137	<800		
DB	73858	CCATCCCAACATGGGATGTATAGACAGACAGACATCTCCCTTAAGTAACTCTA	2406	2464		
QY	2769	GGAGCCCATGGAGCCAGCAAGTAATCTGATCATCTCCAGAGCTTGTGATTTGATA				
DB	73798	GGAGCCCATGGAGCCAGCAAGTAATCTGATCATCTCCAGAGCTTGTGATTTGATA				
QY	2829	AAGCAGGCACTGAGACACCCCTCGTGTCTCTTAACGAAATCTTAAGAGAGAAAT				
DB	73738	AAGCAGGCACTGAGACACCCCTCGTGTCTCTTAACGAAATCTTAAGAGAGAAAT				
QY	2889	ATACGAAAGAAACAGAGAGGTTTCTGTGACACCGCAATTCAATTTGCTGATGGA				
DB	73678	ATACGAAAGAAACAGAGAGGTTTCTGTGACACCGCAATTCAATTTGCTGATGGA				

QY 2949 CTCATTCTAAGGCAAGACATTGAAAATGTAATTCATCTGATACATGTCATGACG 3008  
| | | | |  
Db 73618 CTCATTCTAAGGCAAGACATTGAAAATGTAATTCATCTGATACATGTCATGACG 73559  
| | | | |  
QY 3009 CTCATGTCCTCAACTTAAGGCTGCGGTTTAGCAGCCTGTGAATGAGAGGAGAGAC 3068  
| | | | |  
Db 73558 CTCATGTCCTCAACTTAAGGCTGCGGTTTAGCAGCCTGTGAATGAGAGGAGAGAC 73499  
| | | | |  
QY 3069 CTGAGTCACTGACATGAGGTTTGACAGCAAGCCCTGTGATTCAGAGTGTAAACAGAGCTT 3128  
| | | | |  
Db 73498 CTGAGTCACTGACATGAGGTTTGACAGCAAGCCCTGTGATTCAGAGTGTAAACAGAGCTT 73439  
| | | | |  
QY 3129 GCCCTTCTTGAAGACAAAGTTTCCAAGAGCCTTACCTGAGGTCTCTACTCTCACT 3188  
| | | | |  
Db 73438 GCCCTTCTTGAAGACAAAGTTTCCAAGAGCCTTACCTGAGGTCTCTACTCTCACT 73379  
| | | | |  
QY 3189 GGGGTCCTCCAGAGTGAAGAACGACATGTCCTTTTATTTATTTATTTGGTGGCTG 3248  
| | | | |  
Db 73378 GGGGTCCTCCAGAGTGAAGAACGACATGTCCTTTTATTTATTTATTTGGTGGCTG 73319  
| | | | |  
QY 3249 TGTATTATTAAGATCAAAATGTAATTAACAACCTAGCTCTTTTCACTGACTTATGTAATAAC 3308  
| | | | |  
Db 73318 TGTATTATTAAGATCAAAATGTAATTAACAACCTAGCTCTTTTCACTGACTTATGTAATAAC 73259  
| | | | |  
QY 3309 TCATACTAACTGCTTTGATGCTGCGGTTTGACCTTTACTGACCCCTGATTAACGCTGT 3368  
| | | | |  
Db 73258 TCATACTAACTGCTTTGATGCTGCGGTTTGACCTTTACTGACCCCTGATTAACGCTGT 73199  
| | | | |  
QY 3369 GCGTGTCCCTCCAGAGTGGTGGAAATATTTAACAATCTGTCCAACCAAAAAGAAATGCTGT 3428  
| | | | |  
Db 73198 GCGTGTCCCTCCAGAGTGGTGGAAATATTTAACAATCTGTCCAACCAAAAAGAAATGCTGT 73139  
| | | | |  
QY 3429 GTTTGAGCAGCATTTGACACATATCTGCTTTGATTAAGACATTCCTGATTTCTAGGTCGG 3488  
| | | | |  
Db 73138 GTTTGAGCAGCATTTGACACATATCTGCTTTGATTAAGACATTCCTGATTTCTAGGTCGG 73079  
| | | | |  
QY 3489 TTCGTGTATATCCCATTTGTGGAATTCATCTGTAATCCCATTTGCTTATATGTCCTAGCAA 3548  
| | | | |  
Db 73078 TTCGTGTATATCCCATTTGTGGAATTCATCTGTAATCCCATTTGCTTATATGTCCTAGCAA 73019  
| | | | |  
QY 3549 TAAGAGAAATTTCTCAAGTTTCCATGTGCGGTTTCTAGCGCAGCAATACCTTTAGACA 3608  
| | | | |  
Db 73018 TAAGAGAAATTTCTCAAGTTTCCATGTGCGGTTTCTAGCGCAGCAATACCTTTAGACA 72959  
| | | | |  
QY 3609 TTTAAAGAGAAATTTAGAAATATCTCATCTCTTAAATGTTTAAATATATACCAAAC 3668  
| | | | |  
Db 72958 TTTAAAGAGAAATTTAGAAATATCTCATCTCTTAAATGTTTAAATATATACCAAAC 72899  
| | | | |  
QY 3669 AGTGGCCCCCTGCTATTAGTTTCTGTGTCACCTGCAACCTATTACTTGGTAGCTTAAAAA 3728  
| | | | |  
Db 72898 AGTGGCCCCCTGCTATTAGTTTCTGTGTCACCTGCAACCTATTACTTGGTAGCTTAAAAA 72839  
| | | | |  
QY 3729 CAACCATTAAGCTTAATAGTCTGCGGGAATGAAATTCCAAAATGGAATGCTCCGAAATGAAA 3788  
| | | | |  
Db 72838 CAACCATTAAGCTTAATAGTCTGCGGGAATGAAATTCCAAAATGGAATGCTCCGAAATGAAA 72779  
| | | | |  
QY 3789 ATCAAGGTCTCAGCAGAGCTGTGCTCTTCTGAAGGCTCTAAGGAGAAACCGGTTCTCTG 3848  
| | | | |  
Db 72778 ATCAAGGTCTCAGCAGAGCTGTGCTCTTCTGAAGGCTCTAAGGAGAAACCGGTTCTCTG 72719  
| | | | |  
QY 3849 CCAATTTCAAGCTTCTAGAGGCTGTGCTCATTCACAGGCTTCAGTGGCTGTCAAGCTTTT 3908  
| | | | |  
Db 72718 CCAATTTCAAGCTTCTAGAGGCTGTGCTCATTCACAGGCTTCAGTGGCTGTCAAGCTTTT 72659  
| | | | |  
QY 3909 CTCACATGSCATACATGTAACCTGSCCTCCCACTTCCCTCTTTGACTTAACAAGCCCA 3968  
| | | | |  
Db 72658 CTCACATGSCATACATGTAACCTGSCCTCCCACTTCCCTCTTTGACTTAACAAGCCCA 72599  
| | | | |  
QY 3969 CCAGGAAGATCCAGATTAATCTCTCCATCTAAAGATCTTTCACTCATCTGGAAGAGCTT 4028  
| | | | |  
Db 72598 CCAGGAAGATCCAGATTAATCTCTCCATCTAAAGATCTTTCACTCATCTGGAAGAGCTT 72539  
| | | | |  
QY 4029 TTGCAATGCAAGACAAATAGCCACAGTGGGATTTAGACCAAGACATCTTTGGGCTGC 4088  
| | | | |

Db 72538 TTGCAATGCAAGACAAATAGCCACAGTGGGATTTAGACCAAGACATCTTTGGGCTGC 72479  
| | | | |  
QY 4089 TGTATTCTGCTTACCAACCTCTCTGCACTGACTCCACAGAGAGGCTCAAAATGA 4148  
| | | | |  
Db 72478 TGTATTCTGCTTACCAACCTCTCTGCACTGACTCCACAGAGAGGCTCAAAATGA 72419  
| | | | |  
QY 4149 TCTGCGCAACAGGATGTTTGTGTTAGCTTGGGACTTACACTT-AAAAAACCCGAC 4207  
| | | | |  
Db 72418 TCTGCGCAACAGGATGTTTGTGTTAGCTTGGGACTTACACTTAAAAAAACCCGAC 72359  
| | | | |  
QY 4208 ATCAAGAAATCTGGCAATGCGGGCTCAATCTCACTACCAACAATGCTGAGGT 4267  
| | | | |  
Db 72358 ATCAAGAAATCTGGCAATGCGGGCTCAATCTCACTACCAACAATGCTGAGGT 72299  
| | | | |  
QY 4268 GGGCACACAGCTGTGCTTTAGAGGGGTGTCCACTTCAACAGGTACCAACAGCCCACT 4327  
| | | | |  
Db 72298 GGGCACACAGCTGTGCTTTAGAGGGGTGTCCACTTCAACAGGTACCAACAGCCCACT 72239  
| | | | |  
QY 4328 AGCCCTATCACTTCCCACAATGAGGCTAAGTGTGTTTCTACTGATCAATGCCCTGC 4387  
| | | | |  
Db 72238 AGCCCTATCACTTCCCACAATGAGGCTAAGTGTGTTTCTACTGATCAATGCCCTGC 72179  
| | | | |  
QY 4388 AGTTGCAATTAATGTAATGAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 4447  
| | | | |  
Db 72178 AGTTGCAATTAATGTAATGAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 72119  
| | | | |  
QY 4448 ACCATGAGACCAATGTGCTCACTAACCTTTTCTTTTCTTTTCTTTTCTTTTCTT 4507  
| | | | |  
Db 72118 ACCATGAGACCAATGTGCTCACTAACCTTTTCTTTTCTTTTCTTTTCTTTTCTT 72059  
| | | | |  
QY 4508 TTTTCTTTTATGTAAGACAGAGATCTATCTGTTGCTTGAAGCTGAGTGAAGTGAAG 4567  
| | | | |  
Db 72058 TTTTCTTTTATGTAAGACAGAGATCTATCTGTTGCTTGAAGCTGAGTGAAGTGAAG 71999  
| | | | |  
QY 4568 ATCTGCGCTCACTGAACTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4627  
| | | | |  
Db 71998 ATCTGCGCTCACTGAACTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 71939  
| | | | |  
QY 4628 CAAATGAGTGGATCATCTGCAACAAACCAACATGCCAGCTAATTTTGTATTTTGTAG 4687  
| | | | |  
Db 71938 CAAATGAGTGGATCATCTGCAACAAACCAACATGCCAGCTAATTTTGTATTTTGTAG 71879  
| | | | |  
QY 4688 AGACAGGTTTCAACATGTTGCGCAGGCTGTCTCAACTCTCTGAGCTCAAGCAATCTCTC 4747  
| | | | |  
Db 71878 AGACAGGTTTCAACATGTTGCGCAGGCTGTCTCAACTCTCTGAGCTCAAGCAATCTCTC 71819  
| | | | |  
QY 4748 CTGCTCGGCTCCCAAGTGTGAGGATTTACAGATGTGAGCCACCGCATCTCCAGCCCAAC 4807  
| | | | |  
Db 71818 CTGCTCGGCTCCCAAGTGTGAGGATTTACAGATGTGAGCCACCGCATCTCCAGCCCAAC 71759  
| | | | |  
QY 4808 CCTCATTTATPACAAATTAACCTGCGCAGCTGAACCTGTGACTTTTCTCTCAACCCCTGCT 4867  
| | | | |  
Db 71758 CCTCATTTATPACAAATTAACCTGCGCAGCTGAACCTGTGACTTTTCTCTCAACCCCTGCT 71699  
| | | | |  
QY 4868 CTGATCTGGAAGAGAGGATTTATGTTATGCTTGTGACAGACAGTCCCAAGTTCAATAT 4927  
| | | | |  
Db 71698 CTGATCTGGAAGAGAGGATTTATGTTATGCTTGTGACAGACAGTCCCAAGTTCAATAT 71639  
| | | | |  
QY 4928 TCTGCGGCAAAAACCTTCTTCAAAAATTAATGTAATCTTCTGTAATGTAATGTAATGTAAT 4987  
| | | | |  
Db 71638 TCTGCGGCAAAAACCTTCTTCAAAAATTAATGTAATCTTCTGTAATGTAATGTAATGTAAT 71579  
| | | | |  
QY 4988 TTGGAATGCAACCGCTCAACTTGTTCACATGGAATTAATGAAGAAATTTTATATGCTC 5047  
| | | | |  
Db 71578 TTGGAATGCAACCGCTCAACTTGTTCACATGGAATTAATGAAGAAATTTTATATGCTC 71519  
| | | | |  
QY 5048 CTAAATGCGGTGTACTGCAAGACCTTCTGAACACTTTCAGAGATPAGATATTTAAGTC 5107  
| | | | |  
Db 71518 CTAAATGCGGTGTACTGCAAGACCTTCTGAACACTTTCAGAGATPAGATATTTAAGTC 71459  
| | | | |  
QY 5108 ATGCCCTTGGCGTGTGCTATGAGCACTTTCTTCTGAAGTCTGCTGCTCCAGTGA 5167  
| | | | |

```

Db 71458 ATGCGCTTGGCGTTCCTAATGACCTTTCCTCTCTGAAGTGTGTTCTTCGACATGA 71399
Oy 5168 CCGTTCGCTTGTGAGCCGAGATGCTGACCTGTCAATAAAGGCGCAAGAGGAGGCTGCGGC 5227
Db 71398 CCGTTCGCTTGTGAGCCGAGATGCTGACCTGTCAATAAAGGCGCAAGAGGAGGCTGCGGC 71339
Oy 5228 TTCCTTCCCTCACTGAAGAGCCCTTATTTGAATTCATCTGTGTGAGCCCTAGCCCTCCAT 5287
Db 71338 TTCCTTCCCTCACTGAAGAGCCCTTATTTGAATTCATCTGTGTGAGCCCTAGCCCTCCAT 71279
Oy 5288 TCTGCAATTCGCCCAACCTCCAGCCCTTCCAGAGAGACTAGTGCCCTGCTGATTCAC 5347
Db 71278 TCTGCAATTCGCCCAACCTCCAGCCCTTCCAGAGAGACTAGTGCCCTGCTGATTCAC 71239
Oy 5348 CCAAGTGGAGATTGGCTCTTCTTAGAGCTGCTACTTGTCAACATGACCGACATCACTGTT 5407
Db 71218 CCAAGTGGAGATTGGCTCTTCTTAGAGCTGCTACTTGTCAACATGACCGACATCACTGTT 71159
Oy 5408 GCGTTCAGAGACACACGCGGCTTCTTCACTGAGGCTCAAACTCTGAGACA 5467
Db 71158 GCGTTCAGAGACACACGCGGCTTCTTCACTGAGGCTCAAACTCTGAGACA 71099
Oy 5468 GTTGTGCTCTCTGAGACCACTTCTTCTGAGCTGTGCTCTGAGAGGCGCCAGCTT 5527
Db 71098 GTTGTGCTCTCTGAGACCACTTCTTCTGAGCTGTGCTCTGAGAGGCGCCAGCTT 71039
Oy 5528 GAGGAAACCTGGCTCTTCTTAAAGCCAGGCGCCCTTCAATTAACATTTGAGGT 5587
Db 71038 GAGGAAACCTGGCTCTTCTTAAAGCCAGGCGCCCTTCAATTAACATTTGAGGT 70979
Oy 5588 CACTGGAACAAGTGAAGTCCATTTGTTGAAGCTTACTGATGACAGCCCACTGCTCATC 5647
Db 70978 CACTGGAACAAGTGAAGTCCATTTGTTGAAGCTTACTGATGACAGCCCACTGCTCATC 70919
Oy 5648 CACGTGTCTGTGCACTGCTTCAAGAGAGCCAGCGCATGCAAGCTGCTCTTAATGCTG 5707
Db 70918 CACGTGTCTGTGCACTGCTTCAAGAGAGCCAGCGCATGCAAGCTGCTCTTAATGCTG 70859
Oy 5708 TGGTATTGCAAGAGAGGAAAGTCTCAAGAGAGTCACTGGGAGCAAGCAAGAGCC 5767
Db 70858 TGGTATTGCAAGAGAGGAAAGTCTCAAGAGAGTCACTGGGAGCAAGCAAGAGCC 70799
Oy 5768 ACCGAGACATGCGCTTGTAAAGTTAGACAGACTGTGTGTGTGATCTGCACTGCTTCA 5827
Db 70798 ACCGAGACATGCGCTTGTAAAGTTAGACAGACTGTGTGTGTGATCTGCACTGCTTCA 70739
Oy 5828 TGAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5887
Db 70738 TGAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 70679
Oy 5888 TTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 5947
Db 70678 TTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 70619
Oy 5948 TCCCTGCTTCACTGAGAGGCGCCGCTTCCAGATTGTTGATGATGATGATGATGATG 6007
Db 70618 TCCCTGCTTCACTGAGAGGCGCCGCTTCCAGATTGTTGATGATGATGATGATGATG 70559
Oy 6008 GCTCAGCAGATGCTGTCTTAATTTGTGATGATGATGATGATGATGATGATGATGATG 6067
Db 70558 GCTCAGCAGATGCTGTCTTAATTTGTGATGATGATGATGATGATGATGATGATGATG 70499
Oy 6068 TTTCTTCTCTTCACTTGAATGCGGCTGCTGTGTGATGATGATGATGATGATGATGATG 6127
Db 70498 TTTCTTCTCTTCACTTGAATGCGGCTGCTGTGTGATGATGATGATGATGATGATGATG 70439
Oy 6128 AAAATAATAGTCTTGAAGTCTGCA 6153
Db 70438 AAAATAATAGTCTTGAAGTCTGCA 70413

```

RESULT 8  
AC104165/c

```

LOCUS AC104165 200956 bp DNA linear PRI 27-JUN-2002
DEFINITION Homo sapiens chromosome 3 clone RP11-68104, complete sequence.
ACCESSION AC104165 AC068625
VERSION AC104165.2 GI:21617715
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 200956)
Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,
Saenphimmachak, C., Phelps, K.A., Buckley, D., Kibukawa, M., Raymond, C.
and Haugen, E.D.
Direct Submission
Unpublished
2 (bases 1 to 200956)
Kaul, R.K., Olson, M.V., Raymond, C. and Haugen, E.D.
Direct Submission
Submitted (05-DEC-2001) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
3 (bases 1 to 200956)
Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,
Saenphimmachak, C., Phelps, K.A., Buckley, D., Kibukawa, M., Raymond, C.
and Haugen, E.D.
Direct Submission
Submitted (27-JUN-2002) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
On Jun 27, 2002 this sequence version replaced gi:17352429.
----- Genome Center
Center: University of Washington Genome Center
Center Code: UWGC
Web site: http://www.genome.washington.edu
Contact: uwgchgs@u.washington.edu
Drafting Center: WUSC
----- Project Information
Center project name: chr-3
Center clone name: RP11-68104 (bc0528)
----- Summary Statistics
Sequencing vector: unknown; 4% of reads
Sequencing vector: plasmid; 56% of reads
Chemistry: Dye-terminator BT; 80% of reads
Chemistry: Dye-terminator Big Dye; 20% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 200732 bases at least Q40
Consensus quality: 200947 bases at least Q20
Insert size: 200956; sum-of-contigs
Quality coverage: 10.4x in Q20 bases; sum-of-contigs
-----
Overlapping Sequences:
5': RP11-348P10 (UWGC:bc0377) AC124045
3': RP11-578F5 (UWGC:bc0477) AC105902, 95352-bp overlap
-----
Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp.
Base-pair quality values are not generally visible from the
GenBank flat file format but are available as part
of this entry's ASN.1 file.
-----
This sequence was finished as follows unless otherwise noted:
all regions were either double-stranded or sequenced with an
alternate chemistry or covered by high quality data (i.e., Phred
quality >= 30); an attempt was made to resolve all sequencing
problems, such as compressions and repeats; all regions were
covered by at least one plasmid subclone or more than one M13
subclone; and the assembly was confirmed by restriction digest.
-----
Sequence Validation:
This sequence has been validated by Multiple Complete Digest

```







Db 179683 GGGGGCTGTCTCTCCCTCCCAACCAATATGCTCCAGGGCCCAACTGCAAGTTGGC 179624  
Qy 2649 CACTAGAGAGCCACCTCTCGCTCCCTCTGAGTCTGAGTGAACCGTACACCTTCTC 2708  
Db 179683 CACTAGAGAGCCACCTCTCGCTCCCTCTGAGTCTGAGTGAACCGTACACCTTCTC 179664  
Qy 2709 CCATCCCAACAAATGGGGATGTAAAGCAGCAAGACACATCTCCCTTAAGTCACTCA 2768  
Db 179563 CCATCCCAACAAATGGGGATGTAAAGCAGCAAGACACATCTCCCTTAAGTCACTCA 179504  
Qy 2769 GAGAGCCATGAGCCAGCAGCAATTAATCTGATCTCAATCCAGCGCTTTGCTGATTCATA 2828  
Db 179503 GAGAGCCATGAGCCAGCAGCAATTAATCTGATCTCAATCCAGCGCTTTGCTGATTCATA 179444  
Qy 2829 AAGCAGGGCACTGAGCACCGCTCGTGTCTTAACCAAGAAATCCTAAGAGAGGAAT 2888  
Db 179443 AAGCAGGGCACTGAGCACCGCTCGTGTCTTAACCAAGAAATCCTAAGAGAGGAAT 179384  
Qy 2889 ATACAGAAAGAACAGCAGAGAGGTTTTCTGACACCGCCAACTTGACATTTGCTCAGTGA 2948  
Db 179383 ATACAGAAAGAACAGCAGAGAGGTTTTCTGACACCGCCAACTTGACATTTGCTCAGTGA 179324  
Qy 2949 CTCAATCTTAAGGCGCAAGACATTTGAATATGATTAATCCATCTGGATACATGATGACG 3008  
Db 179323 CTCAATCTTAAGGCGCAAGACATTTGAATATGATTAATCCATCTGGATACATGATGACG 179264  
Qy 3009 CTCAATGCTCTCAACTTAAGGCGTGGGTTAGCGAGCTGTATAGAGAGAGAGGC 3068  
Db 179263 CTCAATGCTCTCAACTTAAGGCGTGGGTTAGCGAGCTGTATAGAGAGAGAGGC 179204  
Qy 3069 CTGAGTCACTAGACATAGGTTGACAGCAAGCCCTGATTCAGAGTGTAAACAGAGGCTT 3128  
Db 179203 CTGAGTCACTAGACATAGGTTGACAGCAAGCCCTGATTCAGAGTGTAAACAGAGGCTT 179144  
Qy 3129 GCCCTCTTACAGCAACAGTCCCAATCCAGAGGCTTCACTGAGTCTCACTCTCACT 3188  
Db 179143 GCCCTCTTACAGCAACAGTCCCAATCCAGAGGCTTCACTGAGTCTCACTCTCACT 179084  
Qy 3189 GGGGCTCCAGGATGAAAAAGCAATGTCCTTTTATATATTAATTTGGTGGTCCG 3248  
Db 179083 GGGGCTCCAGGATGAAAAAGCAATGTCCTTTTATATATTAATTTGGTGGTCCG 179024  
Qy 3249 TGTATTTTAAGATCAATATGATTAACCACTAGCTCTTTTCACTGATTAATTAAC 3308  
Db 179023 TGTATTTTAAGATCAATATGATTAACCACTAGCTCTTTTCACTGATTAATTAAC 178964  
Qy 3309 TCATACTAACTGTTGGATGCTGGGTTGTGACTTCTAAGCCGCTAGATTAACGTT 3368  
Db 178963 TCATACTAACTGTTGGATGCTGGGTTGTGACTTCTAAGCCGCTAGATTAACGTT 178904  
Qy 3369 GCGTGTCCCGAGGTGGTGGGAATTAATTAACAATCTGTCCAAACGAAAAAGATGTGT 3428  
Db 178903 GCGTGTCCCGAGGTGGTGGGAATTAATTAACAATCTGTCCAAACGAAAAAGATGTGT 178844  
Qy 3429 GTTTGAGCAGCATTAACATATCTGCTTTGATTAAGAGACTTCTGATTTCTAGTGGG 3488  
Db 178843 GTTTGAGCAGCATTAACATATCTGCTTTGATTAAGAGACTTCTGATTTCTAGTGGG 178784  
Qy 3489 TTCGTGTTATCCATGTGGAATTCATCTTGAATCCCATGTCTATAGTCTTAGCA 3548  
Db 178783 TTCGTGTTATCCATGTGGAATTCATCTTGAATCCCATGTCTATAGTCTTAGCA 178724  
Qy 3549 TTAAGGAATTTCCCAAGTTTCCATGTCGGGTTCTCTTACGTCAGCAATCTTTGACA 3608  
Db 178723 TTAAGGAATTTCCCAAGTTTCCATGTCGGGTTCTCTTACGTCAGCAATCTTTGACA 178664  
Qy 3609 TTTAAGGAATTTAGAGAAATTTCTCATCTCTTAAATTTAAATATATACCAAC 3668  
Db 178663 TTTAAGGAATTTAGAGAAATTTCTCATCTCTTAAATTTAAATATATACCAAC 178604  
Qy 3669 AGTGAGCCCTGCTGATTAAGTTTCTGTGGCACTGCAACCCATTAATCTGTAGCTTAAAA 3728  
Db 178603 AGTGAGCCCTGCTGATTAAGTTTCTGTGGCACTGCAACCCATTAATCTGTAGCTTAAAA 178544

Qy 3729 CAACATTAAGCTTAATGCTCTGGGGATCAGAAATTCAAAAATGATGTCCCTGAATGAA 3788  
Db 178543 CAACATTAAGCTTAATGCTCTGGGGATCAGAAATTCAAAAATGATGTCCCTGAATGAA 178484  
Qy 3789 ATCAAGGTGTGACAGAGCTGTGCTCTTCTGAAGGCTCTAGAGAAAGCCGTTCTTG 3848  
Db 178483 ATCAAGGTGTGACAGAGCTGTGCTCTTCTGAAGGCTCTAGAGAAAGCCGTTCTTG 178424  
Qy 3849 CCATTTCAAGCTTCTAGAGGCTGTGCTGATTTCCAGGCTCCAGTGGCTGTGTAACCTTT 3908  
Db 178423 CCATTTCAAGCTTCTAGAGGCTGTGCTGATTTCCAGGCTCCAGTGGCTGTGTAACCTTT 178364  
Qy 3909 CTCAATGAGCATCACTGTGACACTGGCCCTCCCACTTCCCTTTTGACTTAAAGCCCA 3968  
Db 178363 CTCAATGAGCATCACTGTGACACTGGCCCTCCCACTTCCCTTTTGACTTAAAGCCCA 178304  
Qy 3969 CCAGGAAGATCCAGATTAATCTCTCATCTAAGATCTTCAATCTCTGAAAGCCTT 4028  
Db 178303 CCAGGAAGATCCAGATTAATCTCTCATCTAAGATCTTCAATCTCTGAAAGCCTT 178244  
Qy 4029 TTGCAATGCAAGAACAAATAGCCCAAGGTGGGATTTAGAACAGACATCTTTGGGGTTC 4088  
Db 178243 TTGCAATGCAAGAACAAATAGCCCAAGGTGGGATTTAGAACAGACATCTTTGGGGTTC 178184  
Qy 4089 TGTATTTCTGCTTACCAACCTTCTGCACTGACTCCCAAGAGAGGCTCAAAATGA 4148  
Db 178183 TGTATTTCTGCTTACCAACCTTCTGCACTGACTCCCAAGAGAGGCTCAAAATGA 178124  
Qy 4149 TCTGCGGACAGAGATGTTTTTTTGTAGCTGGGACTGTAACTT-AAAAAACCAG 4207  
Db 178123 TCTGCGGACAGAGATGTTTTTTTGTAGCTGGGACTGTAACTT-AAAAAACCAG 178064  
Qy 4208 ATCAAGAAATCTGGGCAATGCTGGGCTCAATCTCACTAGCAACAACTGGCTGAGT 4267  
Db 178063 ATCAAGAAATCTGGGCAATGCTGGGCTCAATCTCACTAGCAACAACTGGCTGAGT 178004  
Qy 4268 GGGCAGCAGCTCTGCTTTTGAAGGGGTGTCCACTTCAACAGGTCAACCAAGCCCACT 4327  
Db 178003 GGGCAGCAGCTCTGCTTTTGAAGGGGTGTCCACTTCAACAGGTCAACCAAGCCCACT 177944  
Qy 4328 AGCCCTTACACTTCCCAATAGAGGCTTAAGTGTGTTTCTTCACTGATCAATGCCCCG 4387  
Db 177943 AGCCCTTACACTTCCCAATAGAGGCTTAAGTGTGTTTCTTCACTGATCAATGCCCCG 177884  
Qy 4388 AGTTGCATTTATTTGAATGAAGAAAGAACTGGGATTAATCTTAATCAGGTGATG 4447  
Db 177883 AGTTGCATTTATTTGAATGAAGAAAGAAAGAACTGGGATTAATCTTAATCAGGTGATG 177824  
Qy 4448 ACCATGAGACCAATGAGTCAATTAACCTTTCTTTTCTTTTCTTTTCTTTTCTT 4507  
Db 177823 ACCATGAGACCAATGAGTCAATTAACCTTTCTTTTCTTTTCTTTTCTTTTCTT 177764  
Qy 4508 TTTTCTTTTAAATGAGACAGGATCTAATCTGTTGCTAGGCTGAGTGAAGTGGCGA 4567  
Db 177763 TTTTCTTTTAAATGAGACAGGATCTAATCTGTTGCTAGGCTGAGTGAAGTGGCGA 177704  
Qy 4568 ATCTGGCTCACTGCAACTGTGCTCTGAGCTCAAGCAATCTTCCACTGAGCTTC 4627  
Db 177703 ATCTGGCTCACTGCAACTGTGCTCTGAGCTCAAGCAATCTTCCACTGAGCTTC 177644  
Qy 4628 CAAATAGCTGGGATCACTGGGCAAAACCAACATGCGAGCTAATTTTGAATTTTGTAG 4687  
Db 177643 CAAATAGCTGGGATCACTGGGCAAAACCAACATGCGAGCTAATTTTGAATTTTGTAG 177584  
Qy 4688 AGACAGGTTTCAACCATGTTGCTGAGCTGATCTCAACTCTCTGAGCTCAAGCAATCTCTC 4747  
Db 177583 AGACAGGTTTCAACCATGTTGCTGAGCTGATCTCAACTCTCTGAGCTCAAGCAATCTCTC 177524  
Qy 4748 CTGCTCGGCTCCCAAGGTGCTGGGATTAAGATGTGAGCCACCGCATCCAGCCCA 4807  
Db 177523 CTGCTCGGCTCCCAAGGTGCTGGGATTAAGATGTGAGCCACCGCATCCAGCCCA 177464

```

QY 4808 CCTCATTTATACCAATTACCTGCGCAAGTACCTGAGCTTTGGTCTTCCACCCCTGCT 4867
DB 177463 CCTCATTTATACCAATTACCTGCGCAAGTACCTGAGCTTTGGTCTTCCACCCCTGCT 177404
QY 4868 CTGATGTGGAAGGAGGAGGATTAATGTTATAGCTGTGACAGACAGTCCCAAGTTCAATAT 4927
DB 177403 CTGATGTGGAAGGAGGAGGATTAATGTTATAGCTGTGACAGACAGTCCCAAGTTCAATAT 177344
QY 4928 TCTGCGGCAAAAATTCTTCAAAAATAATGATCTTCAATGATTAATCAATGATTCAC 4987
DB 177343 TCTGCGGCAAAAATTCTTCAAAAATAATGATCTTCAATGATTAATCAATGATTCAC 177284
QY 4988 TTGGAATGACACCGCTCAACTTTGTTCACTGAGCATTAATGAAGAATTTATAGCTTC 5047
DB 177283 TTGGAATGACACCGCTCAACTTTGTTCACTGAGCATTAATGAAGAATTTATAGCTTC 177224
QY 5048 CTAAATGGGAGTACTGCAAGACCTCTTGAACCTTTCAGAGGATAGGATTTATAGTC 5107
DB 177223 CTAAATGGGAGTACTGCAAGACCTCTTGAACCTTTCAGAGGATAGGATTTATAGTC 177164
QY 5108 ATGCCCCCTGGCGTTGGCTATGACACCTTTCCCTTCAAAAAGTCTGGTCTGCCAGTGA 5167
DB 177163 ATGCCCCCTGGCGTTGGCTATGACACCTTTCCCTTCAAAAAGTCTGGTCTGCCAGTGA 177104
QY 5168 CCTTGGCCTTGTGACCCGAGATGCTGACCTGTCATTAAGGCGCAAGAGGCGTGGCG 5227
DB 177103 CCTTGGCCTTGTGACCCGAGATGCTGACCTGTCATTAAGGCGCAAGAGGCGTGGCG 177044
QY 5228 TTCCCTTCCCTCACTGAAGACCCCTTATTTGAATTCACCTGTTGAGGACCCCTAGCCCTCAT 5287
DB 177043 TTCCCTTCCCTCACTGAAGACCCCTTATTTGAATTCACCTGTTGAGGACCCCTAGCCCTCAT 176984
QY 5288 TCTGCAATTTCCCAACTCTCCAGCCCTTCCCAAGAGGACTAGTGGCTGATTCAC 5347
DB 176983 TCTGCAATTTCCCAACTCTCCAGCCCTTCCCAAGAGGACTAGTGGCTGATTCAC 176924
QY 5348 CCAAGTGGAGTTGGCCTTCTTGAAGCTGAGCTACTTGTCAACATCAACGATCACTGTT 5407
DB 176923 CCAAGTGGAGTTGGCCTTCTTGAAGCTGAGCTACTTGTCAACATCAACGATCACTGTT 176864
QY 5408 GCCTGCAAGGACACACGATGGGCCATTTCTTCACTGAAGGAGGCTCAAACTCTCGAGCA 5467
DB 176863 GCCTGCAAGGACACACGATGGGCCATTTCTTCACTGAAGGAGGCTCAAACTCTCGAGCA 176804
QY 5468 GTTGCTGCTCTGAGACCAAGTATTTCTGAGAGCTGTGCTCAGTGAAGGCGCCAGCT 5527
DB 176803 GTTGCTGCTCTGAGACCAAGTATTTCTGAGAGCTGTGCTCAGTGAAGGCGCCAGCT 176744
QY 5528 GAGGAACCTGCTCTTTTCTTTAAAGCCGAGGCCCACTTACATTAACATTTCCAGGT 5587
DB 176743 GAGGAACCTGCTCTTTTCTTTAAAGCCGAGGCCCACTTACATTAACATTTCCAGGT 176684
QY 5588 CACTGGAACAGTGAAGTGCATTTGTTGAAGCTTCTGATGCGAGGCCACAGCTGCATC 5647
DB 176683 CACTGGAACAGTGAAGTGCATTTGTTGAAGCTTCTGATGCGAGGCCACAGCTGCATC 176624
QY 5648 CAGTGTGCTGCGCATGCTACGAGGAAGGCGAGCGATGAGGATGCTCTAATGCTG 5707
DB 176623 CAGTGTGCTGCGCATGCTACGAGGAAGGCGAGCGATGAGGATGCTCTAATGCTG 176564
QY 5708 TGTCTACTGACAGGAAGGAAAGTCTCAAGGAAGTCACTGAGCAAGCAAGCC 5767
DB 176563 TGTCTACTGACAGGAAGGAAAGTCTCAAGGAAGTCACTGAGCAAGCAAGCC 176504
QY 5768 ACCGGAATGGGCTTGTAAAGTTAGAGACCTGAGTGTGTGATCTGAGTGCCTCAG 5827
DB 176503 ACCGGAATGGGCTTGTAAAGTTAGAGACCTGAGTGTGTGATCTGAGTGCCTCAG 176444
QY 5828 TGAATAATTTATTTATTTGAGATCTTTTGAAGTGGATTTTATTTCAATTTCTGCT 5887
DB 176443 TGAATAATTTATTTATTTGAGATCTTTTGAAGTGGATTTTATTTCAATTTCTGCT 176384
QY 5888 TTAATAATTAACAAATGTACAAAAACAAGTATCAAGCTGTTTAAAGTTCGGCTACTTG 5947

```

```

DB 176383 TTAATAATTAACAAATGTACAAAAACAAGTATCAAGCTGTTTAAAGTTCGGCTACTTG 176324
QY 5948 TCCCTGCTTCACTGAGAGGCCCCGGTTTCCAGTTGTTGACTGTGACAGGCTCAGCATG 6007
DB 176323 TCCCTGCTTCACTGAGAGGCCCCGGTTTCCAGTTGTTGACTGTGACAGGCTCAGCATG 176264
QY 6008 GCTCAGCAGATGCTGTCTTAAATTTGTGATGATGATACAGAAAGCCAGGCTTTGGATACAG 6067
DB 176263 GCTCAGCAGATGCTGTCTTAAATTTGTGATGATGATGATACAGAAAGCCAGGCTTTGGATACAG 176204
QY 6068 TTCTTCCCTTCACTTATTTGATGCGCTGCACTGTGTGAGAGAGATGTTTGTCCGGAATA 6127
DB 176203 TTCTTCCCTTCACTTATTTGATGCGCTGCACTGTGTGAGAGAGATGTTTGTCCGGAATA 176144
QY 6128 AAAATAATAGTCTTGAAGTCTGCCA 6153
DB 176143 AAAATAATAGTCTTGAAGTCTGCCA 176118

```

```

RESULT 9
AK026622 3224 bp mRNA linear PRI 29-SEP-2000
LOCUS AK026622
DEFINITION Homo sapiens CDNA: FLJ22969 fls, clone KAT10759.
ACCESSION AK026622
VERSION AK026622.1 GI:10439514
KEYWORDS oligo capping; fls (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (sites)

```

```

REFERENCE
AUTHORS Watanabe,K., Kumagai,A., Itakura,S., Yamazaki,M., Tashiro,H.,
Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,
Nakamura,Y., Isogai,T. and Sugano,S.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 3224)
Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
Shibahara,T., Tanaka,T. and Nakamura,Y.
Direct Submission

```

```

TITLE
JOURNAL Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure Analysis, Human
Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
Japan (E-mail:cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
COMMENT
NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing: Research Association for Biotechnology; cDNA library
construction, 5'- & 3'-end one pass sequencing: Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).

```

```

FEATURES
source
1..3224
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="KAT10759"
/cell_line="KARO III"
/cell_type="signet-ring cell carcinoma"
/clone_1b="KAT"
/notes="cloning vector pMB18SFL3"
111..2621
/notes="unnamed protein product"
/codon_start=1
/protein_id="BAB15511.1"
/db_xref="GI:10439515"
/translacion="MAGLNCVSIALLGLVLAARLPRAAEPAIALPRESNITVLI
KLGPTLAKPCYIVISKRTITMISIKSGERIVFPSCSPENHVIETOKNIDCSG
PCPREGVLOSTSLPLTINRTFIMDVAAHSIGLEIPSLRLIOIGSGSCPCSVT
HSISGRIDATVRICTPCSNGTVERIKRQEGVRAHLHPWPRPVVSGSIANRSIK
RLCTIESVFEESGATLMSANYPEGFPEDELTWQFVVAHLRASVPLNLSNCR

```

BASE COUNT 789 a 922 c 810 g 703 t  
ORIGIN

Query Match 48.7%; Score 3004; DB 9; Length 3224;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 3204; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

KEERVEYIIPGSTNPEVFKLEDPGPMAGNFILSLGCDODASPEILRLQROVLV  
QHPOENSKIIYVVDLSNERAMSLTIEPRPVQSRKFRVGCCELESRTCSNLTLTSG  
SKHISFLCDLRLTLMANVERKISCTDRKQSRKSYSLQVSDIHLLEHLEPSMKL  
LVPRDLSTLVPAOKLOOHTHEKPCSTFSYLVAALPSODLYGSCPGSIIQIO  
VKONISVTLTVPAPSPROASROGLTVEFIEKEBGEVTLTPDKSRVYLRTPMMDR  
GLPSTISVSNMISVPRDOVACTTFPKESGVCCQTRAFMIIOBERTABEELPSLDED  
VLPRSFHHSFWNISCSPTSGKQDLTFSVLTPTPTVDLTLLAAGVGLLS  
ALGLICCVKKKKKKTKNGPAPVGINGIINTEMPROKPFQGRKDNDSHYAVIEDT  
MYGHLLODSSGSFLOPEVDIYRPFQGTMGVCPPSPPTICRPAFLAKLATEBPPRSP  
PESEBEPYTFSPHNPNGDVSKDIDILSTOPEMPERAP"

173 GCGCGCAGTAGTAGTACCGCAGGCGCAGCTGCGCGCGGCTTGGCGCGC 232  
1 GCGCGCAGTAGTAGTACCGCAGGCGCAGCTGCGCGCGGCTTGGCGCGC 60  
233 GCGCGTCCCAACCGTGTTCCTCCCAACGAGGCGAGGCGTCCGGAAGTACGCGCGC 292  
61 GCGCGTCCCAACCGTGTTCCTCCCAACGAGGCGAGGCGTCCGGAAGTACGCGCGC 120  
293 TGAACCTGCGGGGTCTCTATCGCACTGTAGGGGTCTGCTGCTGGGTGCGCGCGCTGC 352  
121 TGAACCTGCGGGGTCTCTATCGCACTGTAGGGGTCTGCTGCTGGGTGCGCGCGCTGC 180  
353 GCGCGGCGGCGAGAGCTTTGAGATTGCTGTGCGCAGAGAAAGCAATTAAGTTCTCA 412  
181 GCGCGGCGGCGAGAGCTTTGAGATTGCTGTGCGCAGAGAAAGCAATTAAGTTCTCA 240  
413 TAAAGCTGGGAGCCCGACTGCTGCGGAGAAACCTGTATACCTCTCTTAAAGAC 472  
241 TAAAGCTGGGAGCCCGACTGCTGCGGAGAAACCTGTATACCTCTCTTAAAGAC 300  
473 ATATACCATGTGTCTCATCAAGTCTGAGAAAGATAGTCTTTAGCTGCGCAGA 532  
301 ATATACCATGTGTCTCATCAAGTCTGAGAAAGATAGTCTTTAGCTGCGCAGA 360  
533 GTCTGAGAAATCATCTTTGTATAGAGATTCAGAAAAATATGACTGTATGCGGCCAT 592  
361 GTCTGAGAAATCATCTTTGTATAGAGATTCAGAAAAATATGACTGTATGCGGCCAT 420  
593 GTCTTTTGGGAGGTTCAAGCTTCAGCCCTCGACATCGTGTGGCTACCTCAACAGA 652  
421 GTCTTTTGGGAGGTTCAAGCTTCAGCCCTCGACATCGTGTGGCTACCTCAACAGA 480  
653 CTTTCATCTGGGATGTCAAAAGCTCATTAAGAGCATCGTTTAGAGTGTGCTTCATCC 712  
481 CTTTCATCTGGGATGTCAAAAGCTCATTAAGAGCATCGTTTAGAGTGTGCTTCATCC 540  
713 CTCGCTGAGGAGATCGGTCCGGGTGAGAGCTGCGCAGAGATCATCTCATCA 772  
541 CTCGCTGAGGAGATCGGTCCGGGTGAGAGCTGCGCAGAGATCATCTCATCA 600  
773 GCGGCGGAATCATCTGCTGCGCAGATTCGAAACCTTTCGACAGCAATGCGACCTGCT 832  
601 GCGGCGGAATCATCTGCTGCGCAGATTCGAAACCTTTCGACAGCAATGCGACCTGCT 660  
833 CCGGATCAAGATGAGAGAGGAGTGAATGAGCTTACCTCCCATGTTCCACCCCA 892  
661 CCGGATCAAGATGAGAGAGGAGTGAATGAGCTTACCTCCCATGTTCCACCCCA 720  
893 GAAATGTCTCGGCTTCAAGCATTTGCAACCGCTCATCTATAAAGCTGTGECATCATCG 952  
721 GAAATGTCTCGGCTTCAAGCATTTGCAACCGCTCATCTATAAAGCTGTGECATCATCG 780  
953 AGTCTGTGTTGAGGGTGAAGGCTCAGCAACCTGTATGTCTGCAACTACCAAGAGGCT 1012  
781 AGTCTGTGTTGAGGGTGAAGGCTCAGCAACCTGTATGTCTGCAACTACCAAGAGGCT 840

1013 TCCTGAGAGTAGCTCATGACGTGCGAGATTGTGCTCTGCGACACTCGCGGCGCAGCG 1072  
841 TCCTGAGAGTAGAGCTCATGACGTGCGAGATTGTGCTCTGCGACACTCGCGGCGCAGCG 900  
1073 TCTCCTCTCAACTTCAACCTCTTCCACTGTGAGAGAGAGAGAGGAGGCTTGAATCT 1132  
901 TCTCCTCTCAACTTCAACCTCTTCCACTGTGAGAGAGAGAGAGGAGGCTTGAATCT 960  
1133 ACATCCCGGGCTCCACACCAACCCCGAGTGTCAAGCTGAGAGAGAGAGAGCTGAGGA 1192  
961 ACATCCCGGGCTCCACACCAACCCCGAGTGTCAAGCTGAGAGAGAGAGAGCTGAGGA 1020  
1193 ACATGCGGAGAACTTCAACCTCTTCTGCAAGAGTGTGACCAAGATGCCAAAGTCAG 1252  
1021 ACATGCGGAGAACTTCAACCTCTTCTGCAAGAGTGTGACCAAGATGCCAAAGTCAG 1080  
1253 GATCTCTCGGCTGAGATTCCAACTTTTGTGTCCAATCCCAAAATGAAGCAATAAAA 1312  
1081 GATCTCTCGGCTGAGATTCCAACTTTTGTGTCCAATCCCAAAATGAAGCAATAAAA 1140  
1313 TCTAGTGTGATGATGATGAGAGAGCATGTCACTACCATGAGAGAGAGAGAGAGAG 1372  
1141 TCTAGTGTGATGATGATGAGAGAGCATGTCACTACCATGAGAGAGAGAGAGAGAG 1200  
1373 TCMAACAGAGCGCAAGTTTGTCCCTGCGCTGTTGTGTCTAGAAATCTGGAACCTGCA 1432  
1201 TCMAACAGAGCGCAAGTTTGTCCCTGCGCTGTTGTGTCTAGAAATCTGGAACCTGCA 1260  
1433 GTAGCAACCTCACTGATCTGCTGCTCCAAACACAAATCTCTTCTTGTATGATC 1492  
1261 GTAGCAACCTCACTGATCTGCTGCTCCAAACACAAATCTCTTCTTGTATGATC 1320  
1493 TGACACGTGTGATGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1552  
1321 TGACACGTGTGATGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380  
1553 AAAGAAATCTTCACTCACTGAGTGCAGTGAATCTTCACTGAGTGCAGTGC 1612  
1381 AAAGAAATCTTCACTCACTGAGTGCAGTGAATCTTCACTGAGTGCAGTGC 1440  
1613 ATGACTTCTCTGGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1672  
1441 ATGACTTCTCTGGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500  
1673 CCGAGACTGAGAGAGATCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1732  
1501 CCGAGACTGAGAGAGATCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560  
1733 TGCCAGTGCATACCCAG 1792  
1561 TGCCAGTGCATACCCAG 1620  
1793 TCAGAGATTCAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1852  
1621 TCAGAGATTCAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680  
1853 TCAG 1912  
1681 TCAG 1740  
1913 AAGGCTTTTCAAGGAG 1972  
1741 AAGGCTTTTCAAGGAG 1800  
1973 GAGAGCGGAG 2032  
1801 GAGAGCGGAG 1860  
2033 AGGTGAGCTGCTGATCTTTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2092  
1861 AGGTGAGCTGCTGATCTTTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920  
2093 CATTCATGATCATCAG 2152

```

Db      1921 CATTATGATCATCCGAGAGCGGAGCCGGCTGAGAGATCTTACCTGGAGAGG 1980
Qy      2153 ATGTGCTCCCAAGCCCAAGCTTCCACATCAAGCTTCTGGGTCAACATCTCTAACTGA 2212
Db      1981 ATGTGCTCCCAAGCCCAAGCTTCCACATCAAGCTTCTGGGTCAACATCTCTAACTGA 2040
Qy      2213 GCGCCCAAGCGGCAAGAGCTAGACCTGCTCTTCTGGGTGACACTTACCCCAAGAGCTG 2272
Db      2041 GCGCCCAAGCGGCAAGAGCTAGACCTGCTCTTCTGGGTGACACTTACCCCAAGAGCTG 2100
Qy      2273 TGGACTTGACTGCTGATCCATCGGAGCGGTGGAGAGTGAAGTCTTACGTGCTGGCCC 2332
Db      2101 TGGACTTGACTGCTGATCCATCGGAGCGGTGGAGAGTGAAGTCTTACGTGCTGGCCC 2160
Qy      2333 TGGGGCTCATCTTGTCTGTGAAAAAAGAAAAAGCAAAAGAGGCGCCGCTG 2392
Db      2161 TGGGGCTCATCTTGTCTGTGAAAAAAGAAAAAGCAAAAGAGGCGCCGCTG 2220
Qy      2393 TGGGTATCTCAATGCGCAATCAATATCTGAGATGCGGAGGCAAGCAAAAGTTTCAGA 2452
Db      2221 TGGGTATCTCAATGCGCAATCAATATCTGAGATGCGGAGGCAAGCAAAAGTTTCAGA 2280
Qy      2453 AAGGGGAAAGCAATGACTCCATGATGATGCAATCAATGAGGACACCATGATATAG 2512
Db      2281 AAGGGGAAAGCAATGACTCCATGATGATGCAATGAGGACACCATGATATAG 2340
Qy      2513 GGCATCTGTCAAGATTCAGAGGCTCTCTCTGAGAGAGAGTGACACCTACCGGC 2572
Db      2341 GGCATCTGTCAAGATTCAGAGGCTCTCTCTGAGAGAGAGTGACACCTACCGGC 2400
Qy      2573 CGTTCCAGGGCAACATGAGGGTCTGTCTCTCTCCCAACCATATCTCCAGGGCCC 2632
Db      2401 CGTTCCAGGGCAACATGAGGGTCTGTCTCTCTCCCAACCATATCTCCAGGGCCC 2460
Qy      2633 CAATGCAAGTTGGGCACTGAGAGGACCTCTGCTCCCTCTGAGTCTGAGAGTG 2692
Db      2461 CAATGCAAGTTGGGCACTGAGAGGACCTCTGCTCCCTCTGAGTCTGAGAGTG 2520
Qy      2693 AACCGTACCTTCTCTCCCAATGAGGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2752
Db      2521 AACCGTACCTTCTCTCCCAATGAGGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2580
Qy      2753 CCTTACTGAACTCAGAGAGCCATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2812
Db      2581 CCTTACTGAACTCAGAGAGCCATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2640
Qy      2813 TTTGCTGATTTCTAATAAGAGGAGCTGAGACACCCGTCCTGTTCTTAACAGAAATC 2872
Db      2641 TTTGCTGATTTCTAATAAGAGGAGCTGAGACACCCGTCCTGTTCTTAACAGAAATC 2700
Qy      2873 CTAAAGAAAGAGAAATATACAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2932
Db      2701 CTAAAGAAAGAGAAATATACAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2760
Qy      2933 CACATTGCTCAGTGAATCTATTCTAAGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2992
Db      2761 CACATTGCTCAGTGAATCTATTCTAAGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2820
Qy      2993 GATTAAGTATGACAGCTCATGCTCTCACTTAAGGCTGTGGCGTTAGCCAGCTGTA 3052
Db      2821 GATTAAGTATGACAGCTCATGCTCTCACTTAAGGCTGTGGCGTTAGCCAGCTGTA 2880
Qy      3053 ATGAGAGAGAGAGAGCTGAGTCACTTACATAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3112
Db      2881 ATGAGAGAGAGAGAGCTGAGTCACTTACATAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2940
Qy      3113 TGTAAACAGAGGCTTGCCTCTTTCAGAGCAACAGATTCCTCAAGAGAGCTTACTGA 3172
Db      2941 TGTAAACAGAGGCTTGCCTCTTTCAGAGCAACAGATTCCTCAAGAGAGCTTACTGA 3000
Qy      3173 GGTCTCTACTCTCAGTGGGGTCCCGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3232

```

```

Db      3001 GGTCTCTACTCTCAGTGGGCTCCAGAGATGAAAAAGCAATGTGCTTTTATTATTAT 3060
Qy      3233 TTAATTTGGTGCTCTGTGTATTATTAAGAGATCAAAATGATTAACCACTAGCTCTTTTAC 3292
Db      3061 TTAATTTGGTGCTCTGTGTATTATTAAGAGATCAAAATGATTAACCACTAGCACTTTTAC 3120
Qy      3293 CTGACTTGTATTAATCTACTACTACTAGTGTGGATGCTGGGTGTGACTTCTACTGAC 3352
Db      3121 CTGACTTGTATTAATCTACTACTACTAGTGTGGATGCTGGGTGTGACTTCTACTGAC 3180
Qy      3353 CGCTAGATAAAGCTGTGCTGCCCCCA 3380
Db      3181 CGCTAGATAAAGCTGTGCTGCCCCCA 3208

RESULT 10
AK026028      2113 bp      mRNA      linear      PRI 29-SEP-2000
LOCUS      AK026028
DEFINITION      Homo sapiens CDNA: FLJ22375 fis, clone HRC06785.
ACCESSION      AK026028
VERSION      AK026028.1 GI:10438737
KEYWORDS      oligo cloning; fis (full insert sequence).
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (sites)
Kawabata,A., Hiki,J., Kobatake,N., Inagaki,H., Ikema,Y.,
Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,
Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 2113)
Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
Shibahara,T., Tanaka,T. and Nakamura,Y.
Direct Submission
Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure Analysis, Human
Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
Japan (E-mail:cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
COMMENT
NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing: Research Association for Biotechnology; cDNA library
construction, 5'- & 3'-end one pass sequencing: Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).
FEATURES
source      1..2113
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="HRC06785"
             /cell_type="primary human renal epithelial cells"
             /clone_11b="HRC"
             /note="Cloning vector pME18FL3"
BASE COUNT      585 a      538 c      470 g      520 t
ORIGIN
Query Match      27.6%; Score 1703; DB 9; Length 2113;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1753; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy      2448 TCAGAAAGGCGAAGAGCAATGATCCCATGATGATGAGTCAATGAGAGACCAATGGT 2507
Db      1 TCAGAAAGGCGAAGAGCAATGATCCCATGATGATGAGTCAATGAGAGACCAATGGT 60
Qy      2508 ATATGGGATCTGTACAGAGATTCCAGGGGCTCTTCTGACGAGAGAGTGAACCTTA 2567
Db      61 ATATGGGATCTGTACAGAGATTCCAGGGGCTCTTCTGACGAGAGTGAACCTTA 120
Qy      2568 CCGGCGTTCCAGGGCACCAATGAGGGGCTGTCTCTCCCTCCACCAACCATATGCTCAG 2627

```

```

Db      121  CGGCGGCTCCAGGGGACCATGGGGTCTGTCCTCCCTCCACCCACCATATGCTCCAG 180
Qy      2658  GGGCCCACTGCAAAAGTTGGGCCACTGAGGAGCCACTCTGCTGCTCCCTCCAGACTCGA 2687
Db      181  GGGCCCACTGCAAAAGTTGGGCCACTGAGGAGCCACTCTGCTGCTCCCTCCAGACTCGA 240
Qy      2688  GAGTGAACCGTACACCTTCTCCCATCCCAACAATGGGAGTGTAGCAGCAAGACAGA 2747
Db      241  GAGTGAACCGTACACCTTCTCCCATCCCAACAATGGGAGTGTAGCAGCAAGACAGA 300
Qy      2748  CATTCCCTTACTGAAACACTCAGAGGCCATGAGCCAGCAAAATTAATTGATCCATTCCA 2807
Db      301  CATTCCCTTACTGAAACACTCAGAGGCCATGAGCCAGCAAAATTAATTGATCCATTCCA 360
Qy      2808  GACGCTTGTGAGTGTCTTAAGAAGAGGCACTGAGACACCCGTCGCTGCTTAACCAAG 2867
Db      361  GACGCTTGTGAGTGTCTTAAGAAGAGGCACTGAGACACCCGTCGCTGCTTAACCAAG 420
Qy      2868  AAATCCTAAAGAGGAATTTATAGAAAGAAACAGCAGAGAGTCTTCTGAGACACCGC 2927
Db      421  AAATCCTAAAGAGGAATTTATAGAAAGAAACAGCAGAGAGTCTTCTGAGACACCGC 480
Qy      2928  AAATCCTAAAGAGGAATTTATAGAAAGAAACAGCAGAGAGTCTTCTGAGAAATTCGA 2987
Db      481  AAATCCTAAAGAGGAATTTATAGAAAGAAACAGCAGAGAGTCTTCTGAGAAATTCGA 540
Qy      2988  ATCTGATATACAGTCAATGACAGTCAATGCTGCTCTCACTTGAAGCTGTGATACCCAG 3047
Db      541  ATCTGATATACAGTCAATGACAGTCAATGCTGCTCTCACTTGAAGCTGTGATACCCAG 600
Qy      3048  CTGTATAGAGAGAGAGAGGCTGAGTCACTAGCATAAGGTTGAGAGAGGCTGAGATT 3107
Db      601  CTGTATAGAGAGAGAGAGGCTGAGTCACTAGCATAAGGTTGAGAGAGGCTGAGATT 660
Qy      3108  CAGAGTGTAAACAGAGGCTTCCCTTTCAGAGCAACAGTTCAATTCAGAGAGGCTTA 3167
Db      661  CAGAGTGTAAACAGAGGCTTCCCTTTCAGAGCAACAGTTCAATTCAGAGAGGCTTA 720
Qy      3168  CCTGAGTGTCCCTACTCTCACTGGGGTCCCAAGATGAAACAGCAATGCTCTTTTAT 3227
Db      721  CCTGAGTGTCCCTACTCTCACTGGGGTCCCAAGATGAAACAGCAATGCTCTTTTAT 780
Qy      3228  ATTATTTATTTGAGTGTCTGTTATTATTAAGAGATCAAAATGTATTAACCACTAGCTCT 3287
Db      781  ATTATTTATTTGAGTGTCTGTTATTATTAAGAGATCAAAATGTATTAACCACTAGCTCT 840
Qy      3288  TTCACCTGACTTATTAATTAATCACTAATCACTGTTTGGATGCTGGGTTGACTTCTA 3347
Db      841  TTCACCTGACTTATTAATTAATCACTAATCACTGTTTGGATGCTGGGTTGACTTCTA 900
Qy      3348  CTGACCCGCTAGATTAACGTCGTGCTGCTCCCAAGTGTGGAGATTAATTAACATCTGTC 3407
Db      901  CTGACCCGCTAGATTAACGTCGTGCTGCTCCCAAGTGTGGAGATTAATTAACATCTGTC 960
Qy      3408  CAACCAAGAAAGATGTGTGTTTGGAGAGATTAATTAATCACTGCTTGTATTAAGAGA 3467
Db      961  CAACCAAGAAAGATGTGTGTTTGGAGAGATTAATTAATCACTGCTTGTATTAAGAGA 1020
Qy      3468  CTTTCTGATTTCTTGAAGTGTGCTGTTGTTATCCCATGTTGAGAAATTCATCTTGAATCC 3527
Db      1021  CTTTCTGATTTCTTGAAGTGTGCTGTTGTTATCCCATGTTGAGAAATTCATCTTGAATCC 1080
Qy      3528  ATTGTCTTATAGTCTTGAAGATTAAGAAATTTCTCAAGTTTCCATGTGCGGTTCTCT 3587
Db      1081  ATTGTCTTATAGTCTTGAAGATTAAGAAATTTCTCAAGTTTCCATGTGCGGTTCTCT 1140
Qy      3588  AGCTGAGCAATATCTTGAATTAAGAAATTTAGAGATTAATTCATCTCTTAATAA 3647
Db      1141  AGCTGAGCAATATCTTGAATTAAGAAATTTAGAGATTAATTCATCTCTTAATAA 1200
Qy      3648  ATGTTTAATATATACAAACAGTGGCCCTGCAATTAATTTTCTGTTGCACTGCAAC 3707

```

```

Db      1201  ATGTTTAATATATATACAAACAGTGGCCCTGCAATTAATTTCTGTTGCCACTGCAAC 1260
Qy      3708  CATTACTGTAGTACTTAAACAAACATTAATAGTCTGAGGATAGAAATTTCCAA 3767
Db      1261  CATTACTGTAGTACTTAAACAAACATTAATAGTCTGAGGATAGAAATTTCCAA 1320
Qy      3768  AATGATGTCCCTGATGAATGAATCAAGGTGTACAGAGACTGTGCTCTTGAAGGCTC 3827
Db      1321  AATGATGTCCCTGATGAATGAATCAAGGTGTACAGAGACTGTGCTCTTGAAGGCTC 1380
Qy      3828  TAGGAGAAAGCCGTTCTTCCATGATTTCAAGCTTTAGAGGCTGCTGATTTCCAGGCT 3887
Db      1381  TAGGAGAAAGCCGTTCTTCCATGATTTCAAGCTTTAGAGGCTGCTGATTTCCAGGCT 1440
Qy      3888  CCAGTGTGCTGATCAAGCTTTTCTACATGAGATCACTGATGACACTGAGCCCTCCACTCC 3947
Db      1441  CCAGTGTGCTGATCAAGCTTTTCTACATGAGATCACTGATGACACTGAGCCCTCCACTCC 1500
Qy      3948  CTCCTTGACTTAACAAAGCCACCAAGAAAGATTCAGGATTAATCTCTCAATTAAGATCCT 4007
Db      1501  CTCCTTGACTTAACAAAGCCACCAAGAAAGATTCAGGATTAATCTCTCAATTAAGATCCT 1560
Qy      4008  TCATATCTCTGAGAGAGCTTTTTCATGATGCAAGACATAGCCACAGGTGGGATTAGG 4067
Db      1561  TCATATCTCTGAGAGAGCTTTTTCATGATGCAAGACATAGCCACAGGTGGGATTAGG 1620
Qy      4068  ACCAGAGATCTTTGGGGTGGCTTATCTGCTACCAACCTTCTGCACTGACTGCC 4127
Db      1621  ACCAGAGATCTTTGGGGTGGCTTATCTGCTACCAACCTTCTGCACTGACTGCC 1680
Qy      4128  ACAGAGAGGCTACAAATATGATCTGCGCACAGGAGTGTGTTTGAAGCTGCGGACTCT 4187
Db      1681  ACAGAGAGGCTACAAATATGATCTGCGCACAGGAGTGTGTTTGAAGCTGCGGACTCT 1740
Qy      4188  AACACTTAATAAAA 4201
Db      1741  AACACTTAATAAAA 1754

```

```

RESULT 11
AR243778      2209 bp      DNA      linear      PAT 20-DEC-2002
LOCUS      AR243778
DEFINITION      Sequence 16 from patent US 6476195.
ACCESSION      AR243778
VERSION      AR243778.1 GI:27291271
KEYWORDS
SOURCE
ORGANISM      Unknown.
REFERENCE
  1 (bases 1 to 2209)
  Komatsoulis,G., Rosen,C.A., Ruben,S.M., Duan,R.D., Moore,P.A.,
  Shi,Y., LaFleur,D.W., Wei,Y.-F., Ni,J., Florence,K.A., Young,P.,
  Brewer,L.A., Soppel,D.R., Endress,G.A., Ebner,R., Olsen,H. and
  Mucenski,M.
  Secreted protein HNF2F20
  Patent: US 6476195-A 16-05-NOV-2002;
  Location/Qualifiers
  source
    /organism="unknown"
BASE COUNT
  508 a      646 c      566 g      477 t      12 others
ORIGIN
Query Match      25.1%; Score 1545; DB 6; Length 2209;
Best Local Similarity 99.4%; Pred.No.0;
Matches 2145; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
Qy      204  AGCTGCGCGGCTTGGGCGCTGAGGCGCGCTGCCACCGTGTTCCTCCACCGAG 263
Db      2  AGCTGCGCGGCTTGGGCGCTGAGGCGCGCTGCCACCGTGTTCCTCCACCGAG 61
Qy      264  GCCGAGGCGTCCCGGAGTCAATGCGCGGCTTAATCTGCGGAGTCTTAATGCACTGTAAG 323
Db      62  GCCGAGGCGTCCCGGAGTCAATGCGCGGCTTAATCTGCGGAGTCTTAATGCACTGTAAG 121

```

QY 324 GGTTCGTGCTGGGTGGCGCGCTGCCGCGGGGCAAGCTTTTGAATGCTCT 383  
DB 122 GGTTCGTGCTGGGTGGCGCGCTGCCGCGGGGCAAGCTTTTGAATGCTCT 181  
QY 384 GCCAGAGAAAGCAATTAAGTTCTCAATTAAGCTGGGGACCCGCACTGTGTGCAAA 443  
DB 182 GCCAGAGAAAGCAATTAAGTTCTCAATTAAGCTGGGGACCCGCACTGTGTGCAAA 241  
QY 444 ACCCTGTATCATGTCATTTCTTAAAGATATATACATGTTGTCCATCAAGCTGTGAGA 503  
DB 242 ACCCTGTATCATGTCATTTCTTAAAGATATATACATGTTGTCCATCAAGCTGTGAGA 301  
QY 504 AAGAATAGTCTTTACCTTTAGCTGCCAGAGTCTTGAAATCACTTTGTATAGATATCA 563  
DB 302 AAGAATAGTCTTTACCTTTAGCTGCCAGAGTCTTGAAATCACTTTGTATAGATATCA 361  
QY 564 GAAATATTTGATGATGTCAGAGCCCAATGCTTTTGGGGAGTTCAGTTCCAGCCCTC 623  
DB 362 GAAATATTTGATGATGTCAGAGCCCAATGCTTTTGGGGAGTTCAGTTCCAGCCCTC 421  
QY 624 GACATGTTGTTGCTTACCTCAACAGAACTTTCACTGGAGTGTCAAAAGCTCATAGAG 683  
DB 422 GACATGTTGTTGCTTACCTCAACAGAACTTTCACTGGAGTGTCAAAAGCTCATAGAG 481  
QY 684 CATGGTTTAAAGCTGCAAGTTTTCATCCCTGCTGAGGCAATCGTTCGGGTGAGAG 743  
DB 482 CATGGTTTAAAGCTGCAAGTTTTCATCCCTGCTGAGGCAATCGTTCGGGTGAGAG 541  
QY 744 CTGCCAGAGCGAGTCACTCACTCAATCAAGCGGCGCAATCAAGTCCAGTTCAGAGAT 803  
DB 542 CTGCCAGAGCGAGTCACTCACTCAATCAAGCGGCGCAATCAAGTTCAGAGAT 601  
QY 804 CGGAACCTTCTGAGCAATGAGCACTGTGTCCCGATCAAGATCAAGAAAGAGTGAAT 863  
DB 602 TGGAACTTCTGAGCAATGAGCACTGTGTCCCGATCAAGATCAAGAAAGAGTGAAT 661  
QY 864 GGCCTTACACCTTCCATGTTCCACCCAGAAATGTCTCCGGCTTCAAGATTGCAAAACG 923  
DB 662 GGCCTTACACCTTCCATGTTCCACCCAGAAATGTCTCCGGCTTCAAGATTGCAAAACG 721  
QY 924 CTCATCATTAATAACGTCGTGATCAATCAAGTGTGTGAGGGAGAGGCTCACCAAC 983  
DB 722 CTCATCATTAATAACGTCGTGATCAATCAAGTGTGTGAGGGAGAGGCTCACCAAC 781  
QY 984 CCTGATGTCTGCAATCACTCAAGAGGCTTCCCTGAGATGAGCTATGAGCTGTGAGTT 1043  
DB 782 CCTGATGTCTGCAATCACTCAAGAGGCTTCCCTGAGATGAGCTATGAGCTGTGAGTT 841  
QY 1044 TGTGTTCTGTGACACCTTGCGGGCCAGCGTCTCTTCTTCAACTTCAACCTTCCAACTG 1103  
DB 842 TGTGTTCTGTGACACCTTGCGGGCCAGCGTCTCTTCTTCAACTTCAACCTTCCAACTG 901  
QY 1104 TTAGAGAGAGAGAGACCGGTTGAATATCAATCCCGGGCTCCACCAACCCCGAGGT 1163  
DB 902 TTAGAGAGAGAGAGACCGGTTGAATATCAATCCCGGGCTCCACCAACCCCGAGGT 961  
QY 1164 GTTCAAGCTGAGAGACAGAGGCTGTGGAACATGAGCGGGAACTTCAACCTTCTGTGCA 1223  
DB 962 GTTCAAGCTGAGAGACAGAGGCTGTGGAACATGAGCGGGAACTTCAACCTTCTGTGCA 1021  
QY 1224 AGGCTGTGACCAAGATGCCCAAGATCCAGAGATCTTCCGGCTGCAAGTTCAAGTTTGGT 1283  
DB 1022 AGGCTGTGACCAAGATGCCCAAGATCCAGAGATCTTCCGGCTGCAAGTTTGGT 1081  
QY 1284 CCAACATCCACAAATGAAGCAATTAATCTACGTGTGTAATCTTGAATATAGCGAGC 1343  
DB 1082 CCAACATCCACAAATGAAGCAATTAATCTACGTGTGTAATCTTGAATATAGCGAGC 1141  
QY 1344 CATGTACTACCATGAGCGCGCCGTCAAAACAGAGCGCAAGTTTGTCTCTGAGCTG 1403  
DB 1142 CATGTACTACCATGAGCGCGCCGTCAAAACAGAGCGCAAGTTTGTCTCTGAGCTG 1201

QY 1404 TTTCTGTGTCTTGAATCTCGGACCTGCAATGCAACTCACCTTGAATCTTGGCTCCA 1463  
DB 1202 TTTCTGTGTCTTGAATCTCGGACCTGCAATGCAACTCACCTTGAATCTTGGCTCCA 1261  
QY 1464 ACACAAATCTCTCTTCTTGTGATGATCTGACAGTCTGTGGAATGGAATGGAATGGA 1523  
DB 1262 ACACAAATCTCTCTTCTTGTGATGATCTGACAGTCTGTGGAATGGAATGGAATGGA 1321  
QY 1524 CATTAAGCTGACAGACCAACCGGTACTGCGCAAGAAATCTTACTCACTCAGGTGCCAG 1583  
DB 1322 CATTAAGCTGACAGACCAACCGGTACTGCGCAAGAAATCTTACTCACTCAGGTGCCAG 1381  
QY 1584 TGAATCTCTCAACCTGCTGTGAGAGTGTATCTTCTCTGAAAGTCTGTGTGCCAA 1643  
DB 1382 TGAATCTCTCAACCTGCTGTGAGAGTGTATCTTCTCTGAAAGTCTGTGTGCCAA 1441  
QY 1644 GGAAGAGCTCAGCTGTGATGCTGTGTCAGCCAGCCAGAAAGCTGACAGCATACACAGAA 1703  
DB 1442 GGAAGAGCTCAGCTGTGATGCTGTGTCAGCCAGCCAGAAAGCTGACAGCATACACAGAA 1501  
QY 1704 GCCCTGCAACACCAAGCTTCACTGCTGAGCTGAGTCCATACCCAGCAGGACCTGTA 1763  
DB 1502 GCCCTGCAACACCAAGCTTCACTGCTGAGCTGAGTCCATACCCAGCAGGACCTGTA 1561  
QY 1764 CTTGCGCTCTCTTCTGCGCGGAGGCTTATCAAGCATCCAGGTGAACAGAACTCTC 1823  
DB 1562 CTTGCGCTCTCTTCTGCGCGGAGGCTTATCAAGCATCCAGGTGAACAGAACTCTC 1621  
QY 1824 GGTGACCTTGTGACCTTGTGCGCCAGCTTCCAAAGAGGCTCCAGGCAAGGTTGTAG 1883  
DB 1622 GGTGACCTTGTGACCTTGTGCGCCAGCTTCCAAAGAGGCTCCAGGCAAGGTTGTAG 1681  
QY 1884 GGTGTCTTTATACCTTATTTTAAAGAGAGAGCGTTTTCAGGTGACCTTGAACAA 1943  
DB 1682 GGTGTCTTTATACCTTATTTTAAAGAGAGAGCGTTTTCAGGTGACCTTGAACAA 1741  
QY 1944 AAGCAAGTCTTCTGAGAGACCCCACTGAGGACCGGGGCTGCTCATCTCTCTGT 2003  
DB 1742 AAGCAAGTCTTCTGAGAGACCCCACTGAGGACCGGGGCTGCTCATCTCTCTGT 1801  
QY 2004 GTTCTGGAACATCAGGTGCGCAGAGACCAAGGTGGCTGCTGACTTCTTTAAGGAGCG 2063  
DB 1802 GTTCTGGAACATCAGGTGCGCAGAGACCAAGGTGGCTGCTGACTTCTTTAAGGAGCG 1861  
QY 2064 GAGCGGCTGTGTCTGACAGACAGGCGGCACTTCAATGATATCCAGAGCAGCGACCG 2123  
DB 1862 GAGCGGCTGTGTCTGACAGACAGGCGGCACTTCAATGATATCCAGAGCAGCGACCG 1921  
QY 2124 GGTGAGAGATCTTCAAGCTGTGAGAGATGTCTCCCAAGCCAAAGCTTCCACATCA 2183  
DB 1922 GGTGAGAGATCTTCAAGCTGTGAGAGATGTCTCCCAAGCCAAAGCTTCCACATCA 1981  
QY 2184 CAGCTTCTGAGTCAATCTCACTCACTGAGGCCCAAGCGGCAAGCAGTGTGACTGTGCT 2243  
DB 1982 CAGCTTCTGAGTCAATCTTAAATTTAAATGAGCCCAAGCGGCAAGCAGTGTGACTGTGCT 2041  
QY 2244 CTTTCTGAGTCACTTACCCCAAGAGCTGTGACTGTGACTGTGACTGTGACTGTGACTGTG 2303  
DB 2042 CTTTCTGAGTCACTTACCCCAAGAGCTGTGACTGTGACTGTGACTGTGACTGTGACTGTG 2101  
QY 2304 GGAAGGTGAGTCTTAACTGTGTCTGTGCGCTCGGGCTCATCATTTTGTGTGTAAAAA 2360  
DB 2102 GGAAGGTGAGTCTTAACTGTGTCTGTGCGCTCGGGCTCATCATTTTGTGTGTAAAAA 2158

RESULT 12  
AK026187 1241 bp mRNA linear PRI 29-SEP-2000  
LOCUS AK026187  
DEFINITION Homo sapiens cDNA: FLJ22534 fls, clone HRC13020.  
ACCESSION AK026187  
VERSION AK026187.1 GI:10438956  
KEYWORDS oligo cloning; fls (full insert sequence).  
SOURCE Homo sapiens (human)



ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases)  
AUTHORS Kawabata, A., Hkaji, T., Kobatake, N., Inagaki, H., Ikema, Y.,  
Okamoto, S., Okitani, R., Ota, T., Suzuki, Y., Odayashi, M., Nishi, T.,  
Shibahara, T., Tanaka, T., Nakamura, Y., Isogai, T. and Sugano, S.  
TITLE NEDO human cDNA sequencing project  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1241)  
AUTHORS Sugano, S., Suzuki, Y., Ota, T., Odayashi, M., Nishi, T., Isogai, T.,  
Shibahara, T., Tanaka, T. and Nakamura, Y.  
TITLE Direct Submission  
JOURNAL Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,  
University of Tokyo, Laboratory of Genome Structure Analysis, Human  
Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,  
Japan (E-mail:cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,  
Fax:81-3-5449-5416)  
COMMENT NEDO human cDNA sequencing project supported by Ministry of  
International Trade and Industry of Japan; cDNA full insert  
sequencing: Research Association for Biotechnology; cDNA library  
construction: 5'- & 3'-end one pass sequencing: Department of  
Virology and Human Genome Center, Institute of Medical Science,  
University of Tokyo (partly supported by Science and Technology  
Agency).  
FEATURES  
source 1..1241  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="HRC13020"  
/cell\_type="Primary human renal epithelial cells"  
/clone\_1b="HRC"  
/note="cloning vector pME18SFL3"  
65..>1241  
55..>1241  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="BA15318.1"  
/db\_xref="GI:10438957"  
/translation="MAGNENLSLOGDDOASPGILRLQFOVLANHPONESKIYVD  
LSNERAMSLTIEPRPVKOSRKVPDGFVLSERTCSNLTLSGSHKISFLCDUTR  
LMNVEKITSCTDHRVCKORSYSLQVPSLILHPELHDFPSKLLVPMKRLSLVTPA  
QRLQOHTHEKPCNTSPSYLVASAIPISDYFSGFCPGGSIKQIOVKONISVLTPTFA  
SEQOASRQGLTFSFIPYKEGCVFTVETDTSKYVLRPNMDRGLSPSTSVSMISV  
PRDOVACLTFKERSGVGVCOTGRAPMIIOERTABEISLDEDVLPRSPFHHSFMV  
NLSNCSPTSGKQDLDFSYTLTPRYVDLTVILIAVGGSVLLSLGLITICVKKKKK  
"  
BASE COUNT 296 a 380 c 313 g 252 t  
ORIGIN  
Query Match 19.1%; Score 1177; DB 9; Length 1241;  
Beet Local Similarity 99.9%; Pred. No. 0;  
Matches 1227; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1133 ACATCCGGGCTCCACCAACCCCGAGGTGTTCAAGCTGAGAGCAAGCAAGCTGGGA 1192  
Db 3 ACATCCGGGCTCCACCAACCCCGAGGTGTTCAAGCTGAGAGCAAGCAAGCTGGGA 62  
QY 1193 ACATGGCGGGAACTTCAACTCTCTCTGCAAGAGCTGTGACCAAGATGCCAAAGTCCAG 1252  
Db 63 ACATGGCGGGAACTTCAACTCTCTCTGCAAGAGCTGTGACCAAGATGCCAAAGTCCAG 122  
QY 1253 GGATCCTCGGGGTGAGCTTCAAGTTTGGTCCACATCCAAATGAAATCAATTAATA 1312  
Db 123 GGATCCTCGGGGTGAGCTTCAAGTTTGGTCCACATCCAAATGAAATCAATTAATA 182  
QY 1313 TCTAGCTGGTTGACTTGAATTAAGAGCGAGCATGTCACTCAATCGAGCCAGCGCCG 1372  
Db 183 TCTAGCTGGTTGACTTGAATTAAGAGCGAGCATGTCACTCAATCGAGCCAGCGCCG 242  
QY 1373 TCAAACAGAGCCGCAAGTTGTCTCGGTGTTTGTGTGTCTAGATCTCGAGCTGCA 1432  
Db 243 TCAAACAGAGCCGCAAGTTGTCTCGGTGTTTGTGTGTCTAGATCTCGAGCTGCA 302

QY 1433 GTAGAACCTCAACCTGACATCTGGCTCCAAAACAAAATCTCTTCTTGTGTGATC 1492  
Db 303 GTAGAACCTCAACCTGACATCTGGCTCCAAAACAAAATCTCTTCTTGTGTGATC 362  
QY 1493 TGACACGTCTGTGATGAATGTGAAAAAACATTAAGCTGCAGACACCGGTACTGCC 1552  
Db 363 TGACACGTCTGTGATGAATGTGAAAAAACATTAAGCTGCAGACACCGGTACTGCC 422  
QY 1553 AAGGAAATCCATCTACCTCCAGGTGCCAGAGACATCTCCACCTGCTGTGAGCTGC 1612  
Db 423 AAGGAAATCCATCTACCTCCAGGTGCCAGAGACATCTCCACCTGCTGTGAGCTGC 482  
QY 1613 ATGACTTCTCTGGAAGCTGTGTGTGCCAAGAGAGCTGAGCTGAGTGTGCTGAG 1672  
Db 483 ATGACTTCTCTGGAAGCTGTGTGTGCCAAGAGAGCTGAGCTGAGTGTGCTGAG 542  
QY 1673 CCAGAGCTGAGAGACATACACAGAGAGCCCTGCAACACAGCTTGACCTACCTCG 1732  
Db 543 CCAGAGCTGAGAGACATACACAGAGAGCCCTGCAACACAGCTTGACCTACCTCG 602  
QY 1733 TGGCAGGCTTACCCAGAGACCTGTACTTGGCTCTTCTTCCCGGAGGCTCTA 1792  
Db 603 TGGCAGGCTTACCCAGAGACCTGTACTTGGCTCTTCTTCCCGGAGGCTCTA 662  
QY 1793 TCAAGCAGATCCAGGTGAAGAGCAAGACATCTCGGTGACCTTGACCTTGGCCCAAGT 1852  
Db 663 TCAAGCAGATCCAGGTGAAGAGCAAGACATCTCGGTGACCTTGACCTTGGCCCAAGT 722  
QY 1853 TCACAACAAGAGCTTCAGAGAGGCTGTGACGCTCTTATTAATTTCAAGAGG 1912  
Db 723 TCACAACAAGAGCTTCAGAGAGGCTGTGACGCTCTTATTAATTTCAAGAGG 782  
QY 1913 AAGGCTTTTACGCTGACCCCTGACACAAAAGAGTCACTGAGAGACCCCACT 1972  
Db 783 AAGGCTTTTACGCTGACCCCTGACACAAAAGAGTCACTGAGAGACCCCACT 842  
QY 1973 GGGACCGGGGCTGACATCCCTCACTCTGTCTCTGGAACATCAGCGTGGCCAGAGAC 2032  
Db 843 GGGACCGGGGCTGACATCCCTCACTCTGTCTCTGGAACATCAGCGTGGCCAGAGAC 902  
QY 2033 AGGTGGCTGCTGACCTTCTTTTAAAGAGCGAGCGGCTGTGCGACAGAGGCGCG 2092  
Db 903 AGGTGGCTGCTGACCTTCTTTTAAAGAGCGAGCGGCTGTGCGACAGAGGCGCG 962  
QY 2093 CATTATGATCATCAAGAGACGCGGACCCGGGCTGAGAGATCTTCAAGCTGACAGG 2152  
Db 963 CATTATGATCATCAAGAGACGCGGACCCGGGCTGAGAGATCTTCAAGCTGACAGG 1022  
QY 2153 ATGTGCTCCCAAGCCAGCTTCCACATCAAGCTTGGGTGCAACATCTTAATGCA 2212  
Db 1023 ATGTGCTCCCAAGCCAGCTTCCACATCAAGCTTGGGTGCAACATCTTAATGCA 1082  
QY 2213 GCCCAGAGCGGCAAGAGAGTGAAGCTGCTCTTCTCGGTGACATTAACCCCAAGACTG 2272  
Db 1083 GCCCAGAGCGGCAAGAGAGTGAAGCTGCTCTTCTCGGTGACATTAACCCCAAGACTG 1142  
QY 2273 TGGACTTGAATCTCATCTCATCGCAGCGGTGGAGGTGAGTCTTAAGTCTGTCTGCC 2332  
Db 1143 TGGACTTGAATCTCATCTCATCGCAGCGGTGGAGGTGAGTCTTAAGTCTGTCTGCC 1202  
QY 2333 TCGGGCTCATATTTGCTGTGAAAAA 2360  
Db 1203 TCGGGCTCATATTTGCTGTGAAAAA 1230  
RESULT 13  
BC021099 1410 bp mRNA linear PRI 14-MAY-2003  
LOCUS Homo sapiens CUB domain-containing protein 1, transcript variant 2,  
DEFINITION BC021099  
ACCESSION BC021099.1 GI:18088978

KEYWORDS  
SOURCE  
ORGANISM

REFERENCE  
AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
JOURNAL

REMARK  
COMMENT

FEATURES  
source

gene  
/db\_xref="locusid:64866"  
/note="synonyms: FLJ22969, MGC31813"  
/gene="CDCP1"  
/clone\_lib="NIH\_MGC\_14"  
/lab\_host="DH10B-R"  
/note="Vector: pOTB7"  
1. 1410  
/db\_xref="locusid:64866"

MGC.  
Homo sapiens (human)  
Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:  
Mammalia: Eutheria: Primates: Catarrhini: Homidae: Homo.  
1 (bases 1 to 1410)  
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,  
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,  
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,  
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heish, F.,  
Diatchenko, L., Marsina, K., Farmer, A., Rubin, G.M., Hong, L.,  
Stephenson, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,  
Schectz, T.E., Brownstein, M.J., Uedini, T.B., Toshiyuki, S.,  
Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,  
Mckernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,  
Morley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,  
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shvachenko, Y.,  
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,  
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,  
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Small, D.E.,  
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.  
human and mouse cDNA sequences  
Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
2 (bases 1 to 1410)  
Strausberg, R.  
Direct Submission  
Submitted (03-JAN-2002) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC Project URL: <http://mgc.ncl.nih.gov>  
Contact: MGC help desk  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: DCTD/DTF  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Genome Sequence Centre,  
BC Cancer Agency, Vancouver, BC, Canada  
infocbcgsc.bc.ca  
Steven Jones, Jennifer Amano, Ian Bosdet, Yaron Butterfield,  
Susana Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,  
Leticia Heiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo  
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven  
Ness, Pawan Pandoh, Anna-Liisa Prabh, Parvaneh Saedi, Jacqueline  
Schein, Duane Small, Michael Smith, Lorraine Spence, Jeff Stott,  
Michael Thorne, Miranda Tsai, Natasja van den Bosch, Jill Vardy,  
George Yang, Scott Zyderdun, Marco Marra.  
Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>  
Series: IRAL Plate: 40 Row: f Column: 14  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 12383093.  
Location/Qualifiers  
1. 1410  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="MGC:31813 IMAGE:4590554"  
/tissue\_type="Kidney, renal cell adenocarcinoma"  
/clone\_lib="NIH\_MGC\_14"  
/lab\_host="DH10B-R"  
/note="Vector: pOTB7"  
1. 1410  
/gene="CDCP1"  
/note="synonyms: FLJ22969, MGC31813"  
/db\_xref="locusid:64866"

CDS  
136..1167  
/codon\_start=1  
/product="CUB domain-containing protein 1, isoform 2"  
/protein\_id="AAH21099.1"  
/db\_xref="gi:18088979"  
/db\_xref="locusid:64866"  
/translation="MAGLNCVSTALLVLLGARRLRGAEAFIALPRESNTYVLI  
KLGPTLLAKPQYIVISRHITMLSIKGERIVTFPSCSPENHVFIEIQKIDWGS  
PCPEEVLQPSSTSLPTNRFIVDAHKSIGLEQFSIPRLQIGRESCPDQV  
HSISGRIDATVRICTPCSNQVSRIRKQEGVXALHPMFIPRVSGSILNRSIK  
RLCIESVREGSGSTLMSANVPEGFPELMTQFVPAHRAVSFLNPLNSER  
KEERXYIIPGTTNPEVFKLEDRKPMNMGAFNLSLGGCDDADSPGILRLQFQVLY  
OHPQNESSE"  
BASE COUNT 321 a 395 c 370 g 324 t  
Query Match 18.8% Score 1159; DB 9; Length 1410;  
Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;  
Matches 1159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
148 GGGCGGGGCTCGGGCGGTCCCGCGCGAGTGTAGTACCGAGGCGAGCGAGCT 207  
1 GGGCGGGGCTCGGGCGGTCCCGCGCGAGTGTAGTACCGAGGCGAGCGAGCT 60  
208 GGGCGGGGCTCGGGCGGTCCCGCGCGAGTGTAGTACCGAGGCGAGCGAGCT 267  
61 GGGCGGGGCTCGGGCGGTCCCGCGCGAGTGTAGTACCGAGGCGAGCGAGCT 120  
268 AGGCGGTCCCGAGTATGCGCGGGCGGTCCCGAGTGTAGTACCGAGGCGAGCT 327  
121 AGGCGGTCCCGAGTATGCGCGGGCGGTCCCGAGTGTAGTACCGAGGCGAGCT 180  
328 CTGCTGTGGGTGCGCGCGGTCCCGCGCGAGTGTAGTACCGAGGCGAGCT 387  
181 CTGCTGTGGGTGCGCGGTCCCGCGCGAGTGTAGTACCGAGGCGAGCT 240  
388 CGAGAAAGCAATTACATTTCTTAAGCTGGGAGCCCGGACCTCTGCTGGCAAAACC 447  
241 CGAGAAAGCAATTACATTTCTTAAGCTGGGAGCCCGGACCTCTGCTGGCAAAACC 300  
448 TGTTCATGTGATTTCTTAAGCAATTACATTTCTTAAGCTGGGAGCCCGGACCT 507  
301 TGTTCATGTGATTTCTTAAGCAATTACATTTCTTAAGCTGGGAGCCCGGACCT 360  
508 ATATGCTTACCTTACAGTGCAGAGTCTCTGAGAAATCACTTTGTCAATAGATCCGAAA 567  
361 ATATGCTTACCTTACAGTGCAGAGTCTCTGAGAAATCACTTTGTCAATAGATCCGAAA 420  
568 AATATTGACTGTATGTCAAGCCCATGTCTTTGGGAGGTTCAAGCTTCAAGCTTGAACA 627  
421 AATATTGACTGTATGTCAAGCCCATGTCTTTGGGAGGTTCAAGCTTCAAGCTTGAACA 480  
628 TGTGTTGGCTTACCTTACAGTGCAGAGTCTCTGAGAAATCACTTTGTCAATAGATCCGAAA 687  
481 TGTGTTGGCTTACCTTACAGTGCAGAGTCTCTGAGAAATCACTTTGTCAATAGATCCGAAA 540  
688 GGTTTAGAGCTCAGTTTTCATCCCTCGCTGAGGAGAGTCTCGGCTGAGAGCTGC 747  
541 GGTTTAGAGCTCAGTTTTCATCCCTCGCTGAGGAGAGTCTCGGCTGAGAGCTGC 600  
748 CCAGAGGAGTCACTCACTCCATCAAGGGCGGAATGATCCACCGTGTCAAGATCCGA 807  
601 CCAGAGGAGTCACTCACTCCATCAAGGGCGGAATGATCCACCGTGTCAAGATCCGA 660  
808 ACCCTTGTGAGAAAGTCACTGTGCTCCGGAGTCAAGATCCAGAAAGGAGTGAAGAGGCC 867  
661 ACCCTTGTGAGAAAGTCACTGTGCTCCGGAGTCAAGATCCAGAAAGGAGTGAAGAGGCC 720  
868 TTACACCTCCCATGTGTCAACCCAGAAATGTCTCCGCTTCAAGATCCAGAAAGGCCCTCA 927  
721 TTACACCTCCCATGTGTCAACCCAGAAATGTCTCCGCTTCAAGATCCAGAAAGGCCCTCA 780  
928 TCTATTAAGAGTGTGTGATCATGAGTGTGTGAGAGGTGAAGGCTTCAGCAACCTTG 987



Db 761 TCTATAAAACGTCGTGATCATCGAGTCGTGTTGGAGGTGAAGGCTCAGCAACCCCTG 840  
Qy 988 ATGTCTGCCCACTACCCAGAAAGGCTTCCCTGAGAGATGAGCTCATGACGCGGAGTTTGC 1047  
Db 841 ATGTCTGCCCACTACCCAGAAAGGCTTCCCTGAGAGATGAGCTCATGAGCGAGTTTGC 900  
Qy 1048 GTTCTGTGACACCTGTGGGCGCAGCGTCTCTTCTTCACTTCAACTCTTCAACTGTGAG 1107  
Db 901 GTTCTGTGACACCTGTGGGCGCAGCGTCTCTTCTTCACTTCAACTCTTCAACTGTGAG 960  
Qy 1108 AGGAAGAGAGAGAGCGGTTGAATATCTATCCCGGCTCCACCAACCCCGAGGTGTC 1167  
Db 961 AGGAAGAGAGAGCGGTTGAATATCTATCCCGGCTCCACCAACCCCGAGGTGTC 1020  
Qy 1168 AAGCTGAGAGACAGAGCGCTGGGAAATGAGGCGGGAATTCACTCTCTGCAAGGCT 1227  
Db 1021 AAGCTGAGAGACAGAGCGCTGGGAAATGAGGCGGGAATTCACTCTCTGCAAGGCT 1080  
Qy 1228 TGTGACCAAGATGCCCAAGTCCAGGATCTCCGCTGCAAGTCCAAAGTTTGTGCA 1287  
Db 1081 TGTGACCAAGATGCCCAAGTCCAGGATCTCCGCTGCAAGTCCAAAGTTTGTGCA 1140  
Qy 1288 CATCCACAAATGAAAGCA 1306  
Db 1141 CATCCACAAATGAAAGCA 1159  
RESULT 14  
AA464028 1376 bp DNA linear PAT 16-JUL-2002  
LOCUS AA464028  
DEFINITION Sequence 161 from Patent WO0140466.  
ACCESSION AA464028  
VERSION AA464028.1 GI:21899031  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE  
AUTHORS 1 Baker, K.P., Bernstein, M., DeForge, L., Desnovers, L., Filvaroff, E.,  
Gao, W.O., Gerlstein, M.E., Goddard, A., Godowski, P.J., Gurney, A.L.,  
Sherwood, S., Smith, V., Stewart, T.A., Tumas, D., Watanabe, C.K.,  
Wood, W.L. and Zhang, Z.  
TITLE Secreted and transmembrane polypeptides and nucleic acids encoding  
same  
JOURNAL Patent: WO 0140466-A 161 07-JUN-2001;  
Genentech Inc. (US)  
FEATURES  
source 1. 1376  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
BASE COUNT 311 a 384 c 352 g 329 t  
ORIGIN  
Query Match 18.1%; Score 1117; DB 6; Length 1376;  
Beef Local Similarity 100.0%; Freq. No. 0;  
Matches 1117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 190 CCAGGCGGAGCGGAGCTGCGCGGCTTGGGCGCTGCGGCGCGCTCCCAACCGTGG 249  
Db 1 CCAGGCGGAGCGGAGCTGCGCGGCTTGGGCGCTGCGGCGCGCTCCCAACCGTGG 60  
Qy 250 TTTTCCCAACCGGAGCGGCTGCGGAGTCAATGCGGCGCTTGAATGCGGAGTCTCT 309  
Db 61 TTTTCCCAACCGGAGCGGAGTCAATGCGGAGTCAATGCGGAGTCTCT 120  
Qy 310 ATCGGACGCTGAGGCTTCTGCTGCTGGGTGGGCGCGCTCCCGCGGCGGAGAAAGT 369  
Db 121 ATCGGACGCTGAGGCTTCTGCTGCTGGGTGGGCGCGCTCCCGCGGCGGAGAAAGT 180  
Qy 370 TTGAGATTGCTCTGCGACGAGAAAGCAATTAAGTTCTATAAAGCTGGGAGCCCG 429

Db 181 TTGAGATTGCTCTGCGACGAGAAAGCAATTAAGTTCTATAAAGCTGGGAGCCCG 240  
Qy 430 ACTGCTGGGAGAAACCCGTTTACATGTCATTTCTTAAAGCATATACATGTTGCC 489  
Db 241 ACTGCTGGGAGAAACCCGTTTACATGTCATTTCTTAAAGCATATACATGTTGCC 300  
Qy 490 ATCAAGCTGAGAAAGATAGTCTTACCTTACCTGCGAGTCTGAGATCACTT 549  
Db 301 ATCAAGCTGAGAAAGATAGTCTTACCTTACCTGCGAGTCTGAGATCACTT 360  
Qy 550 GTCATAGAGATCCAGAAATATTTAGCTGATGTCAGGCGCATGTCCTTTGGGAGTT 609  
Db 361 GTCATAGAGATCCAGAAATATTTAGCTGATGTCAGGCGCATGTCCTTTGGGAGTT 420  
Qy 610 CAGCTTCAGCCCTGACATGCTGTTGCTACCTCCCAACAGAACTTTCATCTGAGATGC 669  
Db 421 CAGCTTCAGCCCTGACATGCTGTTGCTACCTCCCAACAGAACTTTCATCTGAGATGC 480  
Qy 670 AAAGCTATTAAGAGCATGCTTTAGAGCTGAGTTTCCATCCCTGCGCTGAGAGATC 729  
Db 481 AAAGCTATTAAGAGCATGCTTTAGAGCTGAGTTTCCATCCCTGCGCTGAGAGATC 540  
Qy 730 GGTCCGGGTGAGAGCTGCGCAGACGAGTCACTCATCTCATAGCGGCGGAATGATGCC 789  
Db 541 GGTCCGGGTGAGAGCTGCGCAGACGAGTCACTCATCTCATAGCGGCGGAATGATGCC 600  
Qy 790 ACCGTGTCAAGATGAGAAACCTTCTGACAGCAATGGCACTGTGTCCTGGATCAAGTGC 849  
Db 601 ACCGTGTCAAGATGAGAAACCTTCTGACAGCAATGGCACTGTGTCCTGGATCAAGTGC 660  
Qy 850 GAAGGAGTGAATAATGAGCTTACCTCCATGAGTTCCACCCAGAAATGTCCTCGGCTTC 909  
Db 661 GAAGGAGTGAATAATGAGCTTACCTCCATGAGTTCCACCCAGAAATGTCCTCGGCTTC 720  
Qy 910 AGCATTGAACACCGCTCATCTATAAAGCTGTGTCATCATGAGTCTGTGTTAGAGGT 969  
Db 721 AGCATTGAACACCGCTCATCTATAAAGCTGTGTCATCATGAGTCTGTGTTAGAGGT 780  
Qy 970 GAAGGCTCAGCAACCTGATGTCGCCAATCCAGAAAGCTTCCCTGAGATGAGCTC 1029  
Db 781 GAAGGCTCAGCAACCTGATGTCGCCAATCCAGAAAGCTTCCCTGAGATGAGCTC 840  
Qy 1030 ATGACGTGGAGTTGTGTCCTGTCAGACACCGGCGGCGAGGTCTCTCTCAACTTC 1089  
Db 841 ATGACGTGGAGTTGTGTCCTGTCAGACACCGGCGGCGAGGTCTCTCTCAACTTC 900  
Qy 1090 AACCTCTCAACTGTGAGAGAGAGAGGCGGTTGAATCTACATCCCGGCTCCACC 1149  
Db 901 AACCTCTCAACTGTGAGAGAGAGAGGCGGTTGAATCTACATCCCGGCTCCACC 960  
Qy 1150 ACCAACCCCGAGGTGTTCAAGCTGAGAGCAAGACGCTTGGAGCATGCGGAGAACTTC 1209  
Db 961 ACCAACCCCGAGGTGTTCAAGCTGAGAGCAAGACGCTTGGAGAAATGCGGAGAACTTC 1020  
Qy 1210 AACCTCTCTGCAAGGCTGAGAGAGAGATGCCAAAGTCCAGGAGTCTCCGCTGAG 1269  
Db 1021 AACCTCTCTGCAAGGCTGAGAGAGATGCCAAAGTCCAGGAGTCTCCGCTGAG 1080  
Qy 1270 TTCCAAGTTTGTGTCCAATCCATCCAAATGAAAGCA 1306  
Db 1081 TTCCAAGTTTGTGTCCAATCCATCCAAATGAAAGCA 1117  
RESULT 15  
AK026329 976 bp mRNA linear PRI 29-SEP-2000  
LOCUS AK026329  
DEFINITION Homo sapiens cDNA: FLJ22676 fis, clone HS110569.  
ACCESSION AK026329  
VERSION AK026329.1 GI:10439158  
KEYWORDS oligo capping; fis (full insert sequence).  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2004, 03:53:40 ; Search time 11576 Seconds  
(without alignments)  
12939.572 Million cell updates/sec

Title: US-09-899-569A-3

Perfect score: 6163  
Sequence: 1 ccaacgcgcgaatggggagc.....agcttcgcaaaaaa 6163

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 22781392 seqs, 1215238056 residues

Word size: 0

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

EST:  
1: em\_eetba:\*  
2: em\_eetbam:\*  
3: em\_eetbu:\*  
4: em\_eetbu:\*  
5: em\_eetbu:\*  
6: em\_eetbu:\*  
7: em\_eetbu:\*  
8: em\_eetbu:\*  
9: gb\_eet1:\*  
10: gb\_eet2:\*  
11: gb\_eet3:\*  
12: gb\_eet3:\*  
13: gb\_eet4:\*  
14: gb\_eet5:\*  
15: em\_eetfun:\*  
16: em\_eetfun:\*  
17: em\_ges\_hum:\*  
18: em\_ges\_inv:\*  
19: em\_ges\_pln:\*  
20: em\_ges\_vrt:\*  
21: em\_ges\_fun:\*  
22: em\_ges\_mam:\*  
23: em\_ges\_mus:\*  
24: em\_ges\_pro:\*  
25: em\_ges\_rod:\*  
26: em\_ges\_phg:\*  
27: em\_ges\_vrl:\*  
28: gb\_ges1:\*  
29: gb\_ges2:\*

.Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	826	13.4	9 AB033417	AB033417 AB033417
2	826	13.4	9 AB063026	AB063026 EST392 hu
3	766	12.4	13 B0951236	B0951236 AGENCOURT
4	735	11.9	13 B0962997	B0962997 AGENCOURT

5	728	11.8	1008	13 B0691792	B0691792 AGENCOURT
6	720	11.7	868	13 B0681577	B0681577 AGENCOURT
7	706	11.5	882	13 B0153170	B0153170 AGENCOURT
8	704	11.4	1150	10 BE877551	BE877551 601486503
9	662	10.7	676	13 B0157152	B0157152 AGENCOURT
10	642	10.4	643	12 B0044552	B0044552 603622093
11	635	10.3	738	14 CA307237	CA307237 UI-H-FY1-
12	629	10.2	851	10 B0419603	B0419603 602451816
13	628	10.2	683	12 B1457122	B1457122 603185367
14	621	10.1	712	12 B0018619	B0018619 UI-H-DH1-
15	612	9.9	871	13 B0889198	B0889198 AGENCOURT
16	611	9.9	611	10 B0403590	B0403590 602419303
17	594	9.6	717	14 CD367582	CD367582 UI-H-FY1-
18	586	9.5	671	9 AL710125	AL710125 DKFZP686A
19	573	9.3	573	12 B0966953	B0966953 1128e07.Y
20	570	9.2	570	12 B0966788	B0966788 1128e09.Y
21	565	9.2	565	12 B0893567	B0893567 1128e09.X
22	562	9.1	562	14 CB157642	CB157642 K-EST0216
23	559	9.1	690	14 CB128689	CB128689 K-EST0178
24	553	9.0	977	13 B0950224	B0950224 AGENCOURT
25	546	8.9	552	9 AW138207	AW138207 UI-H-BI1-
26	545	8.8	863	10 BE877911	BE877911 601489392
27	537	8.7	614	14 CB122885	CB122885 K-EST0171
28	537	8.7	722	9 AL601433	AL601433 DKFZP313H
29	535	8.7	723	14 CA314193	CA314193 UI-CF-FNO
30	530	8.6	530	12 B0840003	B0840003 K-EST0116
31	522	8.5	522	12 B0893690	B0893690 1128e07.X
32	514	8.3	916	13 B0691185	B0691185 AGENCOURT
33	513	8.3	513	14 CB129853	CB129853 K-EST0179
34	513	8.3	843	10 B0499230	B0499230 602545670
35	505	8.2	505	12 B0840669	B0840669 K-EST0117
36	504	8.2	569	12 B0018832	B0018832 UI-H-DH1-
37	503	8.2	621	10 B0119833	B0119833 602352174
38	495	8.0	495	12 B0701518	B0701518 UI-E-E02
39	488	7.9	594	9 AW578426	AW578426 RC1-CT024
40	488	7.9	594	9 AW604363	AW604363 RC1-CT024
41	486	7.9	542	10 BF336602	BF336602 RC1-CT052
42	480	7.8	480	13 BX473477	BX473477 DKFZP686C
43	479	7.8	481	9 A1275971	A1275971 QW06G08.X
44	478	7.8	478	12 B0838647	B0838647 K-EST0115
45	476	7.7	527	12 B0826981	B0826981 K-EST0099

#### ALIGNMENTS

RESULT 1  
LOCUS AB033417 933 bp mRNA linear EST 18-OCT-1999  
DEFINITION AB033417 Homo sapiens skov Homo sapiens cDNA, mRNA sequence.  
ACCESSION AB033417  
VERSION AB033417.1 GI:6062545  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
AUTHORS Zhou,Y., Kato,H. and Wake,N.  
TITLE 1 (bases 1 to 933)  
JOURNAL Biological function unknown  
COMMENT Unpublished  
Contact: Zhou Yong  
Department of Reproductive Physiology and Endocrinology  
Medical Institute of Bioregulation, Kyushu University  
Tsurumihara 4546, Beppu, Oita 874-0838, Japan  
Tel: 0977-27-1660  
Fax: 0977-27-1661  
Email: yzhou@surumi.beppu.kyushu-u.ac.jp.

#### FEATURES

source  
1..933  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"



Db		481	TTGATTCATTCCAGACGGCTTGTGCGATGTTTCATAAAGCAGGGCACTGAGACACCCTGCCG	540
Oy		2855	TGTTCTTAACCAAGAATCTCTTAAGAAGGAATTTATACAAAGAACGCAGAGGTTTT	2914
Db		541	TGTTCTTAACCAAGAATCTCTTAAGAAGGAATTTATACAAAGGAACGCAGAGGTTTT	600
Oy		2915	CCTGGACACGGCCCAACTTCACATTTGCTCAGAGGACTCATTTCTTAAGGGCAGACAATTGAA	2974
Db		601	CCTGGACACGGCCCAACTTCACATTTGCTCAGAGGACTCATTTCTTAAGGGCAGACAATTGAA	660
Oy		2975	ATGATGAATTCGAATCTGATATACAGTCAAGTCAAGCTCATGTGCTCTCAACTTAGAGCTGT	3034
Db		661	ATGATGAATTCGAATCTGATATACAGTCAAGTCAAGCTCATGTGCTCTCAACTTAGAGCTGT	720
Oy		3035	GCGGTTAGCCACAGCCTGTATATAGAGAGAGAGAGGCGCTGATCACTTAGAGGGTTGCAG	3094
Db		721	GCGGTTAGCCACAGCCTGTATATAGAGAGAGAGAGGCGCTGATCACTTAGAGGGTTGCAG	780
Oy		3095	CAAGCCCCGGAATTCAGAGTGTTAAACAGAGCTTGCCCTCTTCAGAGAACAGTTCCAA	3154
Db		781	CGAGCCCTCGAATTCAGAGTGTTAAACAGAGCTTGCCCTCTTCAGAGAACAGTTCCAA	840
Oy		3155	TCCAAAGAGACTCACTGAGATGCCCTACTCTCACTGGG	3191
Db		841	TCCAAAGAGACTCACTGAGATGCCCTACTCTCACTGGG	877
RESULT 3				
BQ951236				
LOCUS				
DEFINITION		BQ951236	984 bp	mRNA linear EST 21-AUG-2002
ACCESSION		AGENCOURT_8817950 NIH_MGC_18 Homo sapiens	CDNA clone IMAGE:6422238	
VERSION		BQ951236		
KEYWORDS		BQ951236.1 GI:22365714		
SOURCE		EST.		
ORGANISM		Homo sapiens (human)		
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE		1 (bases 1 to 984)		
JOURNAL		NIH-MGC http://mgc.nci.nih.gov/.		
COMMENT		National Institutes of Health, Mammalian Gene Collection (MGC)		
		Unpublished		
		Contact: Robert Stransberg, Ph.D.		
		Email: cga@bsf-research.nih.gov		
		Tissue Procurement: DCTD/BTP/Gasdar		
		CDNA Library Preparation: Rubin Laboratory		
		CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)		
		DNA Sequencing by: Agencourt Bioscience Corporation		
		Clone distribution: MGC clone distribution information can be		
		found through the I.M.A.G.E. Consortium/LNLN at:		
		http://image.lnl.gov		
		Plate: LICM2598 row: 0 column: 07		
		High quality sequence start: 6		
		High quality sequence stop: 719.		
FEATURES				
Source		Location/Qualifiers		
		1..984		
		/organism="Homo sapiens"		
		/mol_type="mRNA"		
		/db_xref="taxon:9606"		
		/clone_image="IMAGE:6422238"		
		/tissue_type="large cell carcinoma"		
		/lab_host="DH10B (phage-resistant)"		
		/clone_id="NIH MGC 18"		
		/note="Organ: lung; Vector: pORF7; Site 1: XhoI; Site 2:		
		EcoRI; CDNA made by oligo-dt priming. Directionally cloned		
		into EcoRI/XhoI sites using the following 5' adaptor:		
		GGCACGAG(G). Library constructed by Ling Hong in the		
		laboratory of Gerald M. Rubin (University of California,		
		Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and		
		Superscript II RT (Life Technologies). Note: this is a		
		NIH_MGC Library."		
BASE COUNT		228 a 301 C 247 G 197 T 1 others		

ORIGIN	Query Match	12.4%	Score 766	DB 13	Length 984
Best Local Similarity	100.0%	Pred. No. 4,6e-156			
Matches 766	Conservative 0	Mismatches 0	Indels 0	Gaps 0	
QY	CTGTGTTTGGAGGCTGGAAGGCTCAGCAACCTTGATGTCGCAACTACCGAAGGCTTCC				1015
DB	23 CTGTGTTTGGAGGCTGGAAGGCTCAGCAACCTTGATGTCGCAACTACCGAAGGCTTCC				82
QY	CTGAGATGAGACTCATGACGNGGAGATTGTCGTTCTCTGACACCTGCGGAGCAGCTCT				1075
DB	83 CTGAGATGAGACTCATGACGNGGAGATTGTCGTTCTCTGACACCTGCGGAGCAGCTCT				142
QY	1076 CCTTCCTCAACTTCAACCTCTCCAACTGTGAGAGAGAGAGAGCGGGTTGAATACTACA				1135
DB	143 CCTTCCTCAACTTCAACCTCTCCAACTGTGAGAGAGAGAGAGCGGGTTGAATACTACA				202
QY	1136 TCCCGGGGCTCCACCAACCCCGAGGTGTTCAAGCTGAGAGACAGACCTTGGAAACA				1195
DB	203 TCCCGGGGCTCCACCAACCCCGAGGTGTTCAAGCTGAGAGACAGACCTTGGAAACA				262
QY	1196 TGGCGGGGAACTTCAACCTCTCTGTGAGAGGCTGTGACAAAGATGCCAAAGTCCAGGA				1255
DB	263 TGGCGGGGAACTTCAACCTCTCTGTGAGAGGCTGTGACAAAGATGCCAAAGTCCAGGA				322
QY	1256 TCCTCCGGCTGCAAGTTTGGTTCACATCCACAAATGAAAGCAATAAATCT				1315
DB	323 TCCTCCGGCTGCAAGTTTGGTTCACATCCACAAATGAAAGCAATAAATCT				382
QY	1316 ACGTGTGTTGACTGATGATGAGGAGGACATGTCACTCAACATGAGCCACCGGCGCTCA				1375
DB	383 ACGTGTGTTGACTGATGATGAGGAGGACATGTCACTCAACATGAGCCACCGGCGCTCA				442
QY	1376 AACAGAGCGCGCAATTTGTGTCCTGCGGCTTTCGATGCTAGAAATCTCGAAGCTGAGTA				1435
DB	443 AACAGAGCGCGCAATTTGTGTCCTGCGGCTTTCGATGCTAGAAATCTCGAAGCTGAGTA				502
QY	1436 GCAACCTCACCTGACATCTGGCTCCAAACAACAATCTCTCTTCTTGTGATGATCTGA				1495
DB	503 GCAACCTCACCTGACATCTGGCTCCAAACAACAATCTCTCTTCTTGTGATGATCTGA				562
QY	1496 CACGTCGTGTGATGATGATGAAAAAACATAGCTGACACAGACACCGGTACTGCCAAA				1555
DB	563 CACGTCGTGTGATGATGATGAAAAAACATAGCTGACACAGACACCGGTACTGCCAAA				622
QY	1556 GGAATTCCTACTCACTCCAGGTGCCAGTACATCTCCACCTGCGCTGTGAGCTGATG				1615
DB	623 GGAATTCCTACTCACTCCAGGTGCCAGTACATCTCCACCTGCGCTGTGAGCTGATG				682
QY	1616 ACTTCTCTGGAAGCTGCTGGTGGCCAAAGACAGAGCTCAGCTGTGCTGATGACAGCC				1675
DB	683 ACTTCTCTGGAAGCTGCTGGTGGCCAAAGACAGAGCTCAGCTGTGCTGATGACAGCC				742
QY	1676 AGAAGCTGACAGACATACACACGAGAAAGCCCTTGCAACACAGCTT				1721
DB	743 AGAAGCTGACAGACATACACACGAGAAAGCCCTTGCAACACAGCTT				788
RESULT 4					
BQ962997	928 bp	mRNA	linear	EST 21-AUG-2002	
LOCUS	AGNC00921	8778970	NIH_MGC_42	Homo sapiens cDNA clone IMAGE:6378668	
DEFINITION	5', mRNA	sequence.			
ACCESSION	BQ962997				
VERSION	BQ962997.1	GI:22378463			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
AUTHORS	Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 928)				
	NIH-MGC	http://mgc.nci.nih.gov/			

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-r@mail.nih.gov  
Tissue Procurement: ATCC

CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: L1CM2564 row: g column: 21  
High quality sequence stop: 589.  
Location/Qualifiers

## FEATURES

1..928  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:637866g"  
/issue\_type="epithelioid carcinoma cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 42"  
/note="Organ: pancreas; Vector: pOTB7; Site\_1: XhoI; Site\_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC library." ]

## BASE COUNT

219 a 257 c 238 g 214 t

## ORIGIN

Query Match 11.9%; Score 735; DB 13; Length 928;  
Best Local Similarity 99.9%; Pred. No. 2.3e-149;  
Matches 785; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

308 CTATCGACTGCTGAGGGGTTCTGCTGCTGCGGCGCGCTCCGCGGCGGCGAGAG 367  
1 CTATCGACTGCTGAGGGGTTCTGCTGCTGCGGCGCGCTCCGCGGCGGCGAGAG 60  
368 CTTTGAAGATTGCTTGCACGAGAAAGCAATTACAGTTCTCATAAAGCTGGGAGCC 427  
61 CTTTGAAGATTGCTTGCACGAGAAAGCAATTACAGTTCTCATAAAGCTGGGAGCC 120  
428 CGACTGCTGGGAAACCTGTTACATGTCATTTCTAAAGACATATACATGTTGT 487  
121 CGACTGCTGGGAAACCTGTTACATGTCATTTCTAAAGACATATACATGTTGT 180  
488 CCATCAAGTCTGGAGAAAGATAGTCTTTACCTTTAGTCCAGAGTCTTGAGATCACT 547  
181 CCATCAAGTCTGGAGAAAGATAGTCTTTACCTTTAGTCCAGAGTCTTGAGATCACT 240  
548 TTGTCTAAGATCCAGAAATATTTGACTGTAATGTCAGGCCCATGCTCTTTGGGAGG 607  
241 TTGTCTAAGATCCAGAAATATTTGACTGTAATGTCAGGCCCATGCTCTTTGGGAGG 300  
608 TTCAGTTAGGCTCGAGCATGTTGTCCTACCTCAACCAAGAACTTTCATCTGGGAGT 667  
301 TTCAGTTAGGCTCGAGCATGTTGTCCTACCTCAACCAAGAACTTTCATCTGGGAGT 360  
668 TCAGAGCTAAGAGCATCGGTTTAGAGCTGAGTTTCAATCCCTCGCTGAGGACA 727  
361 TCAGAGCTAAGAGCATCGGTTTAGAGCTGAGTTTCAATCCCTCGCTGAGGACA 420  
728 TCGGTCGGGAGAGAGCTGCCAGAGTCACTCATCTCATCAGGCGCGAATGATG 787  
421 TCGGTCGGGAGAGAGCTGCCAGAGTCACTCATCTCATCAGGCGCGAATGATG 480  
788 CCACCGTGTGAGGATCGGAACCTTCTGAGCAATGGCACTGTGTCGGATCAAGATGC 847  
481 CCACCGTGTGAGGATCGGAACCTTCTGAGCAATGGCACTGTGTCGGATCAAGATGC 540

QY 848 AAGAAGAGTGAATAATGACCTTTACACCTCCCATGTTCCACCCAGAAATGTTCCGGCT 907  
DB 541 AAGAAGAGTGAATAATGACCTTTACACCTCCCATGTTCCACCCAGAAATGTTCCGGCT 600  
QY 908 TCAGATTCGCAACCGCTCATCTATAAAGCGTCTGATCATGAGTCTGTTGAGG 967  
DB 601 TCAGATTCGCAACCGCTCATCTATAAAGCGTCTGATCATGAGTCTGTTGAGG 660  
QY 968 GTGAAGGCTCAGCAACCTGATGTTGTCGCAACTACCCAGAAAGGCTTCCCTGAGATGAGC 1027  
DB 661 GGGAAAGCTCAGCAACCTGATGTTGTCGCAACTACCCAGAAAGGCTTCCCTGAGATGAGC 720  
QY 1028 TCATGACGTGCGAGTTGTTGTTCTGCTGCAACCTGCGGCGCAGCGTCTTCTTCACT 1087  
DB 721 TCATGACGTGCGAGTTGTTGTTCTGCTGCAACCTGCGGCGCAGCGTCTTCTTCACT 780  
QY 1088 TCACAC 1093  
DB 781 TCACAC 786

RESULT 5  
B0691792 1008 bp mRNA linear EST 15-JUL-2002  
LOCUS AGENCOURT\_8034267 NIH\_MGC\_110 Homo sapiens CDNA clone IMAGE:6208223  
DEFINITION 5', mRNA sequence.  
ACCESSION B0691792  
VERSION B0691792.1 GI:21817108  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1008)  
NIH-MGC http://mgi.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-r@mail.nih.gov  
Tissue Procurement: ATCC

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: L1CM2365 row: i column: 24  
High quality sequence start: 27  
High quality sequence stop: 744.  
Location/Qualifiers

## FEATURES

1..1008  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6208223"  
/issue\_type="ductal carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 110"  
/note="Organ: pancreas; Vector: pOTB7; Site\_1: XhoI; Site\_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC library." ]

BASE COUNT 238 a 238 c 224 g 246 t 62. others

Query Match 11.8%; Score 728; DB 13; Length 1008;  
Best Local Similarity 99.9%; Pred. No. 7e-148;  
Matches 778; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	5389	CTGCACATATCCCAACCTCCAGCCCTCTTCCAAAGAGGACTAGGGCCCTCATTTCCAC	5348
Db	69	CTGCACATATCCCAACCTCCAGCCCTCTTCCAAAGAGGACTAGGGCCCTCATTTCCAC	128
QY	5349	CAAGTGGGATTTGGGCTTCTTAAAGCTGGCTACTTGTCAACATCACCGACATCACTGTG	5408
Db	129	CAAGTGGGATTTGGGCTTCTTAAAGCTGGCTACTTGTCAACATCACCGACATCACTGTG	188
QY	5409	CCTCCAAAGGACACCAACCTGGGCCATTTTCCCTTCAACTAGGGGCTCAAACTCTTGACAA	5468
Db	189	CCTCCAAAGGACACCAACCTGGGCCATTTTCCCTTCAACTAGGGGCTCAAACTCTTGACAA	248
QY	5469	TTTGCTGGCTCTGAGACCAAGTATTTCTTGAGAGCTGTGCTCAGTGAAGGGGCCAGCTG	5528
Db	249	TTTGCTGGCTCTGAGACCAAGTATTTCTTGAGAGCTGTGCTCAGTGAAGGGGCCAGCTG	308
QY	5529	AGGAACCTGGCTCTTTTCTTTAAAGCCAGGCCCACTTACATTAACATTTCAAGGTC	5588
Db	309	AGGAACCTGGCTCTTTTCTTTAAAGCCAGGCCCACTTACATTAACATTTCAAGGTC	368
QY	5589	ACTGGAAACAGTGAAGTGCATTTGTGTAAGCTTACTGCAATGCCAGCCCACTGCTCATCC	5648
Db	369	ACTGGAAACAGTGAAGTGCATTTGTGTAAGCTTACTGCAATGCCAGCCCACTGCTCATCC	428
QY	5649	ACGCGTCTGTCATAGCCCTTACAGAGGAAGGCCAGCGCATGCACAGACTGCTCTAATGCTGT	5708
Db	429	ACGCGTCTGTCATAGCCCTTACAGAGGAAGGCCAGCGCATGCACAGACTGCTCTAATGCTGT	488
QY	5709	GGTCATTGTCACAGAAAGGAAAGTCTCAAGAAAGAGTCMACTGGGACAAAGCAAGCCCA	5768
Db	489	GGTCATTGTCACAGAAAGGAAAGTCTCAAGAAAGAGTCMACTGGGACAAAGCAAGCCCA	548
QY	5769	CCGGAACATGGCTTGGTAAAGTTAGCAGACTGGTGTGTGGATCTGCAAGTCTTCACT	5828
Db	549	CCGGAACATGGCTTGGTAAAGTTAGCAGACTGGTGTGTGGATCTGCAAGTCTTCACT	608
QY	5829	GGAAATATATTAATTCATTGACAGATACCTTTTAAAGTGGAATTTATTCATTTCCGTGTG	5888
Db	609	GGAAATATATTAATTCATTGACAGATACCTTTTAAAGTGGAATTTATTCATTTCCGTGTG	668
QY	5889	TAAATTAACAATGTACCAAAACAAACAGATACAGCTGTTTAAAGTCTTCGGCTACTTGT	5948
Db	669	TAAATTAACAATGTACCAAAACAAACAGATACAGCTGTTTAAAGTCTTCGGCTACTTGT	728
QY	5949	CCCTTGTTCAGTAGAGGCCCGCGTTTCCAGTTGTGACTGTGACAGGCTCAGCATGGG	6008
Db	729	CCCTTGTTCAGTAGAGGCCCGCGTTTCCAGTTGTGACTGTGACAGGCTCAGCATGGG	788
QY	6009	CTACGACAGATCTGTCTTAATTTGTGATGTATACGAAGACGAGCTTTGGGATATCAAG	6067
Db	789	CTACGACAGATCTGTCTTAATTTGTGATGTATACGAAGACGAGCTTTGGGATATCAAG	847

RESULT 6	BO687577	668 bp	RNA	linear	EST_15-JUN-2002
LOCUS	BO687577				
DEFINITION	BO687577	AGNCNCURT_8343949	NIH_MGC_110	Homo sapiens cDNA clone IMAGE:6248364	
ACCESSION	BO687577				
VERSION	BO687577.1				
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 868)				
AUTHORS	NIH-MGC <a href="http://imgc.ncl.nih.gov/">http://imgc.ncl.nih.gov/</a> .				
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)				
JOURNAL	Unpublished				
COMMENT	Contact: Robert Struhsberg, Ph.D. Email: <a href="mailto:cgabbs-remail.nih.gov">cgabbs-remail.nih.gov</a> Tissue Procurement: ATCC				

```

CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINTL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LINTL at:
http://image.llnl.gov
Plate: L1CM388 row: b column: 13
High quality sequence stop: 702.
location/Qualifiers
1..868

```

BASE COUNT	214 a	238 c	206 g	210 t
ORIGIN				

```

Query Match          11.7%;   Score 720;   DB 13;   Length 868;
Best Local Similarity 100.0%;   Pred. No. 4,1e-116;
Matches 720;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0

```

Qy	5065	CAAGACCTCTTGAACA	CTTTCTCGAAGATAGGATATTTAACTCATG	CCCTTGACCTTGCC	5122
Db	1	CAAGACCTCTTGAACA	CTTTCTCGAAGATAGGATATTTAACTCATG	CCCTTGACCTTGCC	60
Qy	5125	TATGGCACTTTTCCCTTCTGAAAGTCTGGTTCTTGCCCA	GTGACCTTGACCTTGAGC	5188	
Db	61	TATGGCACTTTTCCCTTCTGAAAGTCTGGTTCTTGCCCA	GTGACCTTGAGC	120	
Qy	5185	CGAGATGCTGAACCTTGATTAAGGGCCAAAGAGAGGCTG	CGGCTTCTTCCCTCATGTAA	5244	
Db	121	CGAGATGCTGAACCTTGATTAAGGGCCAAAGAGAGGCTG	CGGCTTCTTCCCTCATGTAA	180	
Qy	5245	GAGCCCTTATTTGAAATTCATGTTGAGGCCCTGACCCCT	CCATTCCTCGACATTTCCCAAC	5304	
Db	181	GAGCCCTTATTTGAAATTCATGTTGAGGCCCTGACCCCT	CCATTCCTCGACATTTCCCAAC	240	
Qy	5305	CTCCACAGCCCTTTCAGACAGAGACTAGGTGCTTCGAT	TTCCACCMAAGTGGGATTTGGC	5366	
Db	241	CTCCACAGCCCTTTCAGACAGAGACTAGGTGCTTCGAT	TTCCACCMAAGTGGGATTTGGC	300	
Qy	5365	TTCTTTAGGCTGGCTACTTGTCAACATCA	CCGACATCACTGTGCTTGCAAGACACAC	5422	
Db	301	TTCTTTAGGCTGGCTACTTGTCAACATCA	CCGACATCACTGTGCTTGCAAGACACAC	360	
Qy	5425	GTGGCCATTTTCCCTTCAACTGAGGGGCTCAAAAC	CTCGGACAAAGTGTGGCTCCTGAGA	5484	
Db	361	GTGGCCATTTTCCCTTCAACTGAGGGGCTCAAAAC	CTCGGACAAAGTGTGGCTCCTGAGA	420	
Qy	5485	CCAAGTATTTCTCTGAGCTGTGCTCAAGTGAAGGGGGCC	AGCCTGAGNAAACCTTGCTCTT	5544	
Db	421	CCAAGTATTTCTCTGAGCTGTGCTCAAGTGAAGGGGGCC	AGCCTGAGNAAACCTTGCTCTT	480	
Qy	5545	TTCTTTAAAGCCCAAGGCCCACTTAACATTA	AAACATTTCAAGGCTCACTTGAAACATGTAAAG	5604	
Db	481	TTCTTTAAAGCCCAAGGCCCACTTAACATTA	AAACATTTCAAGGCTCACTTGAAACATGTAAAG	540	
Qy	5605	TGCACTTTGTTGAAGCTTACTGACATGTCAGAGCCCA	CTGTCATTCACACGAGTGTCTGGCATGAC	5666	
Db	541	TGCACTTTGTTGAAGCTTACTGACATGTCAGAGCCCA	CTGTCATTCACACGAGTGTCTGGCATGAC	600	
Qy	5665	CTACGAGAAAGCCACGAGCATGACAGACTGGTCTCTA	TATGCTGTGTTGATTTGACACAGAG	5724	

DB 601 CTACGAGAAAGCCACCGATGCGAGATGCTCTTAATGCTGTGCTCAATGACAGAG 660  
QY 5725 GGAAAGCTCTCAAGAAAGTCAACTGAGCAAGCAAGCAAGCCAGCCGAGATGGCTTGG 5784  
DB 661 GGAAAGCTCTCAAGAAAGTCAACTGAGCAAGCAAGCAAGCCAGCCGAGATGGCTTGG 720  
RESULT 7  
BUI53170 882 bp mRNA linear EST 03-SEP-2002  
LOCUS AGENCOURT 7937392 NIH\_MGC\_70 Homo sapiens cDNA clone IMAGE:6026162  
DEFINITION 5' mRNA sequence.  
ACCESSION BUI53170 GI:22666702  
VERSION BUI53170.1  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE NIH-MGC http://mgi.nci.nih.gov/  
AUTHORS 1 (bases 1 to 882)  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaabs-r@mail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM13239 row: h column: 03  
High quality sequence stop: 645.  
Location/Qualifiers  
1..882  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6026162"  
/issue\_type="epithelioid carcinoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_id="NIH\_MGC\_70"  
/note="Organ: pancreas; Vector: pCMV-Sport6; Site: 1: NCI;  
Site 2: Sall; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 1.1 kb. Library constructed by Life  
Technologies."

BASE COUNT 205 a 258 c 184 g 235 t  
ORIGIN  
Query Match 11.5%; Score 706; DB 13; Length 882;  
Best Local Similarity 99.8%; Pred. No. 4.3e-143;  
Matches 806; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4630 AATAGCTGGATCACTGGCAAAACCAACCAATGCGCCAGTAAATTTGTATTTTGTAGAG 4689  
DB 3 AATAGCTGGATCACTGGCAAAACCAACCAATGCGCCAGTAAATTTGTATTTTGTAGAG 62  
QY 4690 AAGAGGTTTACACATGTCGCCAGGCTGTCTCAACCTCTGGGCTCAGCAATCTCTCT 4749  
DB 63 AAGAGGTTTACACATGTCGCCAGGCTGTCTCAACCTCTGGGCTCAGCAATCTCTCT 122  
QY 4750 GCGTCGCGCTCCCAAAAGTCTGGATTAAGATGTAGCCACGATCCAGCCCAAC 4809  
DB 123 GCGTCGCGCTCCCAAAAGTCTGGATTAAGATGTAGCCACGATCCAGCCCAAC 182  
QY 4810 CTCATTATACCAATTAACCTGCGCACTGAGACTTTTGCTTCTCAACCTGCTCT 4869  
DB 183 CTCATTATACCAATTAACCTGCGCACTGAGACTTTTGCTTCTCAACCTGCTCT 242  
QY 4870 GATCTGGAAGAGAGGATTAATGTTTACTTGTACGACAGTCCCAAGTCAATTTTC 4929

DB 243 GATCTGGAAGAGAGGATTAATGTTTACTTGTACGACAGTCCCAAGTCAATTTTC 302  
QY 4930 TCGGCAAAAACCTCTTCAAAAATAATGACTTCAATTTCAATGAATTCACCTT 4989  
DB 303 TCGGCAAAAACCTCTTCAAAAATAATGACTTCAATTTCAATGAATTCACCTT 362  
QY 4990 GGAATGACCGCCCTCACTTGTTCATGATTAAGAAAGAAATTTTATAGTCTCT 5049  
DB 363 GGAATGACCGCCCTCACTTGTTCATGATTAAGAAAGAAATTTTATAGTCTCT 422  
QY 5050 AAATGGCTGTACTGCAAGACCTCTTGAACACTTTCAGAGATAGATTAATTAATCAT 5109  
DB 423 AAATGGCTGTACTGCAAGACCTCTTGAACACTTTCAGAGATAGATTAATTAATCAT 482  
QY 5110 GCCCTTGGGCTTCTTATGACACCTTCTCTTGAAGATCTGTTCTGCCAGTACC 5169  
DB 483 GCCCTTGGGCTTCTTATGACACCTTCTCTTGAAGATCTGTTCTGCCAGTACC 542  
QY 5170 CTGGGCTTTGTAGAGCCGAGATGCTGTAACCTTGATTAAGGCGCAAGAGGCTGGGCTT 5229  
DB 543 CTGGGCTTTGTAGAGCCGAGATGCTGTAACCTTGATTAAGGCGCAAGAGGCTGGGCTT 602  
QY 5230 CCTTCCCTCACTGAGAGCCCTTATTTGAATTCATCTGTGAGGCTTACCTTCATTC 5289  
DB 603 CCTTCCCTCACTGAGAGCCCTTATTTGAATTCATCTGTGAGGCTTACCTTCATTC 662  
QY 5290 TCGACATTTCCCAACCTCCAGCCCTTTCAGAGAGACTAGTGTCCCTGATTCACCC 5349  
DB 663 TCGACATTTCCCAACCTCCAGCCCTTTCAGAGAGACTAGTGTCCCTGATTCACCC 722  
QY 5350 AAGGTGGATTTGGCTTCTTCTTATGAGCTGTCTTCAACATCAGCATCATCTGTTC 5409  
DB 723 AAGGTGGATTTGGCTTCTTCTTATGAGCTGTCTTCAACATCAGCATCATCTGTTC 782  
QY 5410 CTGCAAGACACACAGTGGCCATTTTC 5437  
DB 783 CTGCAAGACACACAGTGGCCATTTTC 810  
RESULT 8  
BE877551 1150 bp mRNA linear EST 20-OCT-2000  
LOCUS BE877551  
DEFINITION 601486503P1 NIH\_MGC\_69 Homo sapiens cDNA clone IMAGE:3888768 5',  
mRNA sequence.  
ACCESSION BE877551  
VERSION BE877551.1 GI:10326327  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE NIH-MGC http://mgi.nci.nih.gov/  
AUTHORS 1 (bases 1 to 1150)  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaabs-r@mail.nih.gov  
Tissue Procurement: DCTD/DRP/Gazdar  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM9669 row: f column: 01  
High quality sequence stop: 724.  
Location/Qualifiers  
1..1150  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3888768"  
/issue\_type="Large cell carcinoma, undifferentiated"





```

Db      541 TTAGCAGACTGGTGTGTGATCTGCAGTGTCTCACTGAAATATTTATTCATTGACG 600
Oy      5851 ATACTTTAGGTGATTTATTTATTTCTGTCGTTAAATTAACAATGATACCAAA 5910
Db      601 ATACTTTTAGGTGATTTATTTATTTCTGTCGTTAAATTAACAATGATACCAAA 660
Oy      5911 AA 5912
Db      661 AA 662

RESULT 10
BM044552      643 bp      mRNA      linear      EST 07-NOV-2001
LOCUS      60362203F1 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:5447542 5',
DEFINITION      mRNA sequence.
ACCESSION      BM044552
VERSION      BM044552.1 GI:16773819
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE      NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL      National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT      Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgephs-remail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLCM1931 row: j column: 23
High quality sequence stop: 639.
Location/Qualifiers
1. 643
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5447542"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/notes="Organ: prostate; Vector: pOT87; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

BASE COUNT      153 a      185 c      127 g      178 t

Query Match      10.4%; Score 642; DB 12; Length 643;
Best Local Similarity 100.0%; Pred. No. 3.3e-129;
Matches 642; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      4529 GATCTATCTGTGCTAGCTGAGTGCAGTGGCGCAATCTCGGCTCACTGCAACCTC 4588
Db      1 GATCTATCTGTGCTAGCTGAGTGCAGTGGCGCAATCTCGGCTCACTGCAACCTC 60
Oy      4589 TGCCTCTGGGCTCAAGCAATTCCTCCACCTCAGCTCCCAATAGCTGGGATGATGCG 4648
Db      61 TGCCTCTGGGCTCAAGCAATTCCTCCACCTCAGCTCCCAATAGCTGGGATGATGCG 120
Oy      4649 ACAACACCATGCGCAGCTAATTTGTATTTTGTAGAGACAGGGTTTACCAATGTTG 4708

```

```

Db      121 ACAACACCATGCGCAGCTAATTTGTATTTTGTAGAGACAGGGTTTACCAATGTTG 180
Oy      4709 CCCAGCTGTGCTCAACCTCTCGGCTCAAGCAATCTCTGCTCGGCTCCAAAGTG 4768
Db      181 CCCAGCTGTGCTCAACCTCTCGGCTCAAGCAATCTCTGCTCGGCTCCAAAGTG 240
Oy      4769 CTGGATTTACAGATGTGAGCCACCGCATCCAGCCCAACCTCATTTATACCAATYAC 4828
Db      241 CTGGATTTACAGATGTGAGCCACCGCATCCAGCCCAACCTCATTTATACCAATYAC 300
Oy      4829 TGCCAGTAACTGTGAGCTTTTGTCTTCTCAACCCCTGCTGATCTGGAAGAGAGGAT 4888
Db      301 TGCCAGTAACTGTGAGCTTTTGTCTTCTCAACCCCTGCTGATCTGGAAGAGAGGAT 360
Oy      4889 TATGTTATAGCTTGTACAGACAGTCCCAAGTCAATTTTCTGCGCAAAAATCTTCTTC 4948
Db      361 TATGTTATAGCTTGTACAGACAGTCCCAAGTCAATTTTCTGCGCAAAAATCTTCTTC 420
Oy      4949 AAAAAATTAATGTAATCTTCAATTTGATTTCAATGAAATTCAGTGAATGACCGCTCAAC 5008
Db      421 AAAAAATTAATGTAATCTTCAATTTGATTTCAATGAAATTCAGTGAATGACCGCTCAAC 480
Oy      5009 TTGTTCACTGAGCATTAATGAAGAATTTTATAGTCTCTTAATGAGCGGTACTGCAAG 5068
Db      481 TTGTTCACTGAGCATTAATGAAGAATTTTATAGTCTCTTAATGAGCGGTACTGCAAG 540
Oy      5069 ACCCTTGAAACATTTCCAGAGATAGATATTTAATGATGATCCCTTGCGCTGCTATG 5128
Db      541 ACCCTTGAAACATTTCCAGAGATAGATATTTAATGATGATCCCTTGCGCTGCTATG 600
Oy      5129 GCACCTTTCCTCTTGAAAGTGTGTTCTGCTGCGCGTACCC 5170
Db      601 GCACCTTTCCTCTTGAAAGTGTGTTCTGCTGCGCGTACCC 642

RESULT 11
CA307237/c      738 bp      mRNA      linear      EST 01-NOV-2002
LOCUS      UI-H-FPI1-bhu-p-08-0-UI-g1 NCI CGAP FPI1 Homo sapiens cDNA clone
DEFINITION      UI-H-FPI1-bhu-p-08-0-UI 3', mRNA sequence.
ACCESSION      CA307237
VERSION      CA307237.1 GI:24470291
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE      NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
JOURNAL      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
COMMENT      Tumor Gene Index
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgephs-remail.nih.gov
Tissue Procurement: Dr. Gary W. Hunninghake, U of I
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
Seq primer: M13 FORWARD
POLYA=yes.
Location/Qualifiers
1. 738
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-FPI1-bhu-p-08-0-UI"
/tissue_type="Aveolar Macrophage"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI_CGAP_FPI1"
/notes="Organ: Lung; Vector: pYT73-Pac (Pharmacia) with a

```

modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP FTL is a normalized cDNA library constructed from a pool of 81 RNA samples from Alveolar Macrophages challenged with different treatments. The library was normalized according to Bonaldi, Lennon and Soars, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GGGCATTGCGG. The tissue was provided by Dr. Gary W. Hummshake of the University of Iowa.

TAG LIB=UI-H-FTL  
TAG TISSUE=Human Lung Alveolar Macrophage  
TAG SEQ=GGCATTGCGG

BASE COUNT 145 a 195 c 229 g 168 t 1 others

ORIGIN

Query Match 10.3%; Score 635; DB 14; Length 738;  
Best Local Similarity 100.0%; Pred. No. 9.7e-128; Indels 0; Gaps 0;  
Matches 635; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1726 TACCTGTCGAGCCAGTCCATACCCAGCCAGCCTTCTGCTCTTCCCGGGA 1785  
648 TACCTGTCGAGCCAGTCCATACCCAGCCAGCCTTCTGCTCTTCCCGGGA 589  
1786 GGGCTCTATCAGCAGATCCAGGTGAAGCAAAATCTCGGTGACCTTGGACCTTTGGC 1845  
588 GGGCTCTATCAGCAGATCCAGGTGAAGCAAAATCTCGGTGACCTTGGACCTTTGGC 529  
1846 CCCAGTTCACAAAGAGGCTCCAGGAGGCTGACGCTGCTTCTTACCTTATTTTC 1905  
528 CCCAGTTCACAAAGAGGCTCCAGGAGGCTGACGCTGCTTCTTACCTTATTTTC 469  
1906 AAAGAGGAAGGGCTTTTCAAGGTGACCCCTGACACAAAAGCAAGGTCTACCTGAGAGC 1965  
468 AAAGAGGAAGGGCTTTTCAAGGTGACCCCTGACACAAAAGCAAGGTCTACCTGAGAGC 409  
1966 CCCAACTGGAGCCGGGGCTGCTCCATCCTTCTGTGTCTGTAACATCAAGCGTGGC 2025  
408 CCCAACTGGAGCCGGGGCTGCTCCATCCTTCTGTGTCTGTAACATCAAGCGTGGC 349  
2026 AAGAGACAGAGTGGCTGCTGCTGCTTCTTAAAGAGGAGGCGGCTGCTGCGCAGAA 2085  
348 AAGAGACAGAGTGGCTGCTGCTGCTTCTTAAAGAGGAGGCGGCTGCTGCGCAGAA 289  
2086 GGGCGGCAATTCATGATCATTCAGAGGAGGAGCCCGGGCTGAGGAGATCTTCAAGCTG 2145  
288 GGGCGGCAATTCATGATCATTCAGAGGAGGAGCCCGGGCTGAGGAGATCTTCAAGCTG 229  
2146 GAGAGAGATGTCTCCCAAGCCAGCTTCCACATCAGAGCTTCTGGGTCAACATCTCT 2205  
228 GAGAGAGATGTCTCCCAAGCCAGCTTCCACATCAGAGCTTCTGGGTCAACATCTCT 169  
2206 AACTGCAAGCCCAAGGCGGCAAGGCTGAGAGCTGCTTCTGCGTGAACATTTACCCA 2265  
168 AACTGCAAGCCCAAGGCGGCAAGGCTGAGAGCTGCTTCTGCGTGAACATTTACCCA 109  
2266 AAGAGCTGAGACTGATGCTGATCCTCATCGAGGCGGTGAGAGTGGAGTCTTACGCTG 2325  
108 AAGAGCTGAGACTGATGCTGATCCTCATCGAGGCGGTGAGAGTGGAGTCTTACGCTG 49  
2326 TGTGCTCTGGGCTCATCATTTGCTGTGTGAAAA 2360  
48 TGTGCTCTGGGCTCATCATTTGCTGTGTGAAAA 14

RESULT 12  
BG419603 851 bp mRNA linear EST 14-MAR-2001  
LOCUS BG419603

DEFINITION 602451816F1 NIH\_MGC\_14 Homo sapiens cDNA clone IMAGE:4590554 5', mRNA sequence.  
ACCESSION BG419603  
VERSION BG419603.1 GI:13326109  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 851)  
NIH-MGC <http://mgc.ncl.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
Tissue Procurement: DCTP/DRP  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed By: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Plate: LHCN1325 row: 0 column: 03  
High quality sequence stop: 699.  
Location/Qualifiers

## FEATURES

source

1..851  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4590554"  
/tissue\_type="renal cell adenocarcinoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 14"  
/note="Organ: kidney; Vector: pOTB1; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 181 a 236 c 240 g 194 t

Query Match 10.2%; Score 629; DB 10; Length 851;  
Best Local Similarity 100.0%; Pred. No. 1.7e-126; Indels 0; Gaps 0;  
Matches 629; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

194 GGGGAGCGCAGCTGCGCGGCTTGGGCGCTGAGGCGCGCGCTCCCAAGCGTCTT 253  
47 GGGGAGCGCAGCTGCGCGGCTTGGGCGCTGAGGCGCGCGCTCCCAAGCGTCTT 106  
254 CCCACCGAGGCGAGGCGTCCCGAGTCAATGCGCGGCTGAACTGCGGGGCTCTATCG 313  
107 CCCACCGAGGCGAGGCGTCCCGAGTCAATGCGCGGCTGAACTGCGGGGCTCTATCG 166  
314 CACTGCTAGGGGTTTGTGCTGCTGCGGTGGCGCGCGCTCCCGCGGCGAGAGCTTTG 373  
167 CACTGCTAGGGGTTTGTGCTGCTGCGGTGGCGCGCGCTCCCGCGGCGAGAGCTTTG 226  
374 AGATTGCTCTGCGCAGAGAAAGCAATTAACAGTTCTATAAGCTGGGAGCCCGAGCTC 433  
227 AGATTGCTCTGCGCAGAGAAAGCAATTAACAGTTCTATAAGCTGGGAGCCCGAGCTC 286  
434 TGTGCGCAAAACCTGTATACATGTCATTTCTAAAAAGCATTAACATGTTGTCATCA 493  
287 TGTGCGCAAAACCTGTATACATGTCATTTCTAAAAAGCATTAACATGTTGTCATCA 346  
494 AGTCTGGAGAAAGATATGCTTTTACCTTAAGTGCAGAGCTGAGAGTCACTTTGCA 553  
347 AGTCTGGAGAAAGATATGCTTTTACCTTAAGTGCAGAGCTGAGAGTCACTTTGCA 406  
554 TAGAGATCCAGAAATAATATGCTATATGTCAGGCCCATGTCCTTTGGGAGGTTCAAGC 613



Seq primer: M13 FORWARD  
POLYA=yes  
FEATURES  
source

Location/Qualifiers  
1..712  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5823821"  
/tissue\_type="Metastatic Chondrosarcoma"  
/dev\_stage="Adult"  
/lab\_host="DH10B (Life Technologies)"  
/clone\_lib="NCI\_CGAP\_DH1"  
/note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP DH1 is a normalized cDNA library containing the following tissue(s): VS-8 Cell line from Metastatic Chondrosarcoma in Lung. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is AGATCATTCG.  
TAG\_LIB=UI-H-DH1  
TAG\_TISSUE=Lung  
TAG\_SEQ=AGATCATTCG"

BASE COUNT 174 a 163 c 192 g 181 t 2 others  
ORIGIN

Query Match 10.1%; Score 621; DB 12; Length 712;  
Best Local Similarity 100.0%; Pred. No. 1e-124;

Matches 621; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

5292 GACATTTCCCAACCTCCAGCCCTTCCAGAGAGACTAGTGCCTGATTCACCA 5351

Db 633 GACATTTCCCAACCTCCAGCCCTTCCAGAGAGACTAGTGCCTGATTCACCA 574

Qy 5332 GGTGGGATTTGGCTTCTTAGGCTGCTACTTGTCACTACCGACATCACTGTTGCTT 5411

Db 573 GGTGGGATTTGGCTTCTTAGGCTGCTACTTGTCACTACCGACATCACTGTTGCTT 514

Qy 5412 GCAAGAGACCACTGGGCTTCTTCACTGAGGGGCTCAAACTCTGGAGCAAGTTG 5471

Db 513 GCAAGAGACCACTGGGCTTCTTCACTGAGGGGCTCAAACTCTGGAGCAAGTTG 454

Qy 5472 CTGGCTCTGAGACCAAGTATTTCTGAGCTGTGCTCAGTGAGAGGGGCCAGCCTGAGG 5531

Db 453 CTGGCTCTGAGACCAAGTATTTCTGAGCTGTGCTCAGTGAGAGGGGCCAGCCTGAGG 394

Qy 5532 AACCTGGCTCTTCTTTAAAGCCAGGCCCTTACATTAACATTTCAAGGCTCACT 5591

Db 393 AACCTGGCTCTTCTTTAAAGCCAGGCCCTTACATTAACATTTCAAGGCTCACT 334

Qy 5592 GGAAGACAGTGAAGTGCATTTGTGAAGCTACTGACAGCCGACCTGCTCACTACAG 5651

Db 333 GGAAGACAGTGAAGTGCATTTGTGAAGCTACTGACAGCCGACCTGCTCACTACAG 274

Qy 5652 TGGTCTGCATGCTTCAAGAGAGCCAGCGCATGAGACTGTCTTAATGCTGTGGT 5711

Db 273 TGGTCTGCATGCTTCAAGAGAGCCAGCGCATGAGACTGTCTTAATGCTGTGGT 214

Qy 5712 CATTGACAGAAAGGAAAGTCTCAAGAGAGTCAACTGGAGCAAGCAAGCCACCG 5771

Db 213 CATTGACAGAAAGGAAAGTCTCAAGAGAGTCAACTGGAGCAAGCAAGCCACCG 154

Qy 5772 GACATGGGCTTGTAAAGTTAGCAGACTGTGTGTGTGAGATCTGCACTGCTTCACTGA 5831

Db 153 GACATGGGCTTGTAAAGTTAGCAGACTGTGTGTGTGAGATCTGCACTGCTTCACTGA 94

Qy 5832 AATAATTATTCATTCAGATTAATTTAGTGAGCATTTTATTCATTTCTGTGCTTAA 5891

Db 93 AATAATTATTCATTCAGATTAATTTAGTGAGCATTTTATTCATTTCTGTGCTTAA 34  
Qy 5892 AATAAAGTATGACCAAAAA 5912  
Db 33 AATAAAGTATGACCAAAAA 13

RESULT 15  
BO889198  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
BO889198 871 bp mRNA linear EST 16-AUG-2002  
AGENCOURT 8072366 NIH\_MGC\_110 Homo sapiens cDNA clone IMAGE:6084315  
5', mRNA sequence.  
BO889198  
BO889198.1 GI:22281212  
EST.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
NIH-MGC http://mgc.nci.nih.gov/  
1 (bases 1 to 871)  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: cga@bbs-remail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.lnl.gov  
Plate: L1CM2315 row: 0 column: 04  
High quality sequence stop: 707.  
Location/Qualifiers  
1..871  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6084315"  
/tissue\_type="ductal carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_110"  
/note="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;  
Site 2: EcoRI; cDNA made by oligo-dT priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCACGAG(G). Library constructed by  
Ling Hong in the laboratory of Gerald M. Rubin (University  
of California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC Library."

FEATURES  
source

BASE COUNT 211 a 247 c 176 g 237 t  
ORIGIN

Query Match 9.9%; Score 612; DB 13; Length 871;  
Best Local Similarity 99.7%; Pred. No. 7.9e-123;

Matches 782; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

3983 GATAATCTCTCCATCTAAGATTCCTTCATCTGAGAGAGCCTTTGCCATGCAAGAC 4042

Db 1 GATAATCTCTCCATCTAAGATTCCTTCATCTGAGAGAGCCTTTGCCATGCAAGAC 60

Qy 4043 AACATAGCCACAGTGGGAGTTAGAGCCAGAGACATCTTTGGGGTGTGTTATTCGCTTA 4102

Db 61 AACATAGCCACAGTGGGAGTTAGAGCCAGAGACATCTTTGGGGTGTGTTATTCGCTTA 120

Qy 4103 CCACACCTTCTGCACTGACTCCACAGGAGAGCTCAAAATGATCTGGCGACAGG 4162

Db 121 CCACACCTTCTGCACTGACTCCACAGGAGAGCTCAAAATGATCTGGCGACAGG 180

Qy 4163 ATGTTTGTAGCTTGGAGACTTAACACTT-AAAAAACCAGATCAGAGATCTGG 4221



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Comphen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2004, 04:02:55 ; Search time 335 Seconds  
(without alignments)  
8120.112 Million cell updates/sec

Title: US-09-899-569a-3

Perfect score: 6163

Sequence: 1 ccaacgcccaatg99agtc.....agctccgcaaaaaaaaaa 6163

Scoring table: OLIGO\_NUC  
Gapop\_60.0, Gapext 60.0

Searched: 569978 seqs, 220691566 residues

Word size: 0

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: Issued Patents, NA:\*

- 1: /cgn2\_6/prodata/1/ina/5A\_COMB.seq:\*
- 2: /cgn2\_6/prodata/1/ina/5B\_COMB.seq:\*
- 3: /cgn2\_6/prodata/1/ina/6A\_COMB.seq:\*
- 4: /cgn2\_6/prodata/1/ina/6B\_COMB.seq:\*
- 5: /cgn2\_6/prodata/1/ina/ECTUS\_COMB.seq:\*
- 6: /cgn2\_6/prodata/1/ina/backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1545	25.1	2209	4	US-09-489-847-16
2	236	3.8	636	3	US-09-328-111-370
3	174	2.8	174	4	US-09-313-294A-7155
4	53	0.9	43069	4	US-09-292-542A-1
5	51	0.8	45546	4	US-09-146-053-6
6	49	0.8	70000	4	US-09-851-896-3
7	45	0.7	2326	4	US-08-994-076-1
8	45	0.7	2418	4	US-09-285-379-1
9	45	0.7	2418	4	US-09-369-912-1
10	45	0.7	14796	3	US-08-975-080-35
11	45	0.7	14796	3	US-09-630-706-10
12	45	0.7	14796	4	US-09-496-694B-3
13	44	0.7	43950	4	US-09-735-934A-3
14	44	0.7	43950	4	US-10-060-332-3
15	44	0.7	176373	3	US-09-128-155-17
16	42	0.7	885	4	US-09-288-143-55
17	42	0.7	1237	4	US-09-904-615-50
18	42	0.7	25464	4	US-09-326-480A-4
19	42	0.7	32654	4	US-09-801-191A-3
20	42	0.7	72604	4	US-09-268-992-7
21	42	0.7	72604	4	US-09-657-474-7
22	41	0.7	162450	4	US-09-345-882-1
23	40	0.6	8220	4	US-09-179-908-3
24	39	0.6	14636	3	US-09-173-914-6
25	39	0.6	40000	4	US-09-780-049-18
26	39	0.6	43950	4	US-09-735-934A-3
27	39	0.6	43950	4	US-10-060-332-3

C	28	39	0.6	62804	4	US-09-800-960-3	Sequence 3, Appl
C	29	39	0.6	99500	4	US-09-798-096-10	Sequence 10, Appl
C	30	39	0.6	174493	4	US-09-804-471A-3	Sequence 3, Appl
C	31	39	0.6	246240	2	US-08-724-394A-20	Sequence 20, Appl
C	32	39	0.6	246240	2	US-08-724-394A-21	Sequence 21, Appl
C	33	39	0.6	246240	2	US-08-724-394A-22	Sequence 22, Appl
C	34	38	0.6	2892	2	US-08-874-186-44	Sequence 44, Appl
C	35	38	0.6	3609	4	US-09-705-299-11	Sequence 11, Appl
C	36	38	0.6	4895	3	US-09-053-866-1	Sequence 1, Appl
C	37	38	0.6	4895	4	US-09-479-130-1	Sequence 1, Appl
C	38	38	0.6	4895	4	US-09-472-130A-1	Sequence 1, Appl
C	39	38	0.6	9377	4	US-09-801-874-3	Sequence 3, Appl
C	40	38	0.6	17327	1	US-07-906-871-15	Sequence 15, Appl
C	41	38	0.6	18596	3	US-09-318-448-11	Sequence 11, Appl
C	42	38	0.6	18597	4	US-09-962-665-8	Sequence 8, Appl
C	43	38	0.6	19011	1	US-08-310-356-36	Sequence 36, Appl
C	44	38	0.6	19557	5	PCT-US92-06300-1	Sequence 1, Appl
C	45	38	0.6	19650	4	US-09-819-989-3	Sequence 3, Appl

## ALIGNMENTS

RESULT 1  
US-09-489-847-16  
; Sequence 16, Application US/09489847  
; Patent No. 6476195  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al  
; TITLE OF INVENTION: 98 Human Secreted Proteins  
; FILE REFERENCE: P2031P1  
; CURRENT APPLICATION NUMBER: US/09/489, 847  
; CURRENT FILING DATE: 2000-01-24  
; EARLIER FILING DATE: 1998-07-29  
; EARLIER FILING DATE: 1998-07-29  
; EARLIER FILING DATE: 1998-07-30  
; EARLIER FILING DATE: 1998-07-30  
; EARLIER FILING DATE: 1998-08-05  
; EARLIER FILING DATE: 1998-08-05  
; EARLIER FILING DATE: 1998-08-12  
; EARLIER FILING DATE: 1998-08-12  
; EARLIER FILING DATE: 1998-08-06  
; EARLIER FILING DATE: 1998-08-06  
; EARLIER FILING DATE: 1998-08-06  
; EARLIER FILING DATE: 1998-08-06  
; NUMBER OF SEQ ID NOS: 376  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 16  
; LENGTH: 2209  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-489-847-16

Query Match 25.1%; Score 1545; DB 4; Length 2209;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 2145; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY	204	AGCTGCGCGGCTTGGGCGCTGCGCGCGCTTCCACCGTGTTCACCGG	263
DB	2	AGCTCGCGCGGCTTGGGCGCTGCGCGCGCTTCCACCGTGTTCACCGG	61
QY	264	GCGGAGGCGTCCCGGAGTCATGCGCGCTGAACTGCGGCGTCTATCGCACTG	323
DB	62	GCGGAGGCGTCCCGGAGTCATGCGCGCTGAACTGCGGCGTCTATCGCACTG	121
QY	334	GCTTCTGCTGCTGCGTGGCGCGCTGCGCGCGCGGAGCTTTGAGATTGCT	383
DB	122	GCTTCTGCTGCTGCGTGGCGCGCTGCGCGCGCGGAGCTTTGAGATTGCT	181
QY	384	GCGGAGGAGCAATTAACATTTCAATTAAGCTGGGAGCCCGGACTTGTGCGCAA	443
DB	182	GCGGAGGAGCAATTAACATTTCAATTAAGCTGGGAGCCCGGACTTGTGCGCAA	241



444 ACCCTGTAATCATGTCATTTCTAAAGACATATACCATGTTGTCCATCATAGCTGGAGA 503  
DB 242 ACCCTGTAATCATGTCATTTCTAAAGACATATACCATGTTGTCCATCATAGCTGGAGA 301  
QY 504 AAGAAATAGCTTAACTTAAAGCTGAGAGCTGAGAAATCACTTGTATAGAGATCA 563  
DB 302 AAGAAATAGCTTAACTTAAAGCTGAGAGCTGAGAAATCACTTGTATAGAGATCA 361  
QY 564 GAAAAATATGACTGTATGTCAGAGCCCATGTCCTTTTGGGAGGTTCACTTCAGCCCTC 623  
DB 362 GAAAAATATGACTGTATGTCAGAGCCCATGTCCTTTTGGGAGGTTCACTTCAGCCCTC 421  
QY 624 GACATGTTGTTGCTCACTTCAAGAACTTTCATCTGGAGTTCATAGCTCATAGAG 683  
DB 422 GACATGTTGTTGCTCACTTCAAGAACTTTCATCTGGAGTTCATAGCTCATAGAG 481  
QY 684 CATCGGTTTAAAGCTGAGCTGATTTCCATCCCTGAGAGAGATGAGTCCGGGTGAGAG 743  
DB 482 CATCGGTTTAAAGCTGAGCTGATTTCCATCCCTGAGAGAGATGAGTCCGGGTGAGAG 541  
QY 744 CTGCCGAGACGAGAGTCACTCATCTCATGAGCGCGCATGATGCGACCGTGGTCAGAT 803  
DB 542 CTGCCGAGACGAGAGTCACTCATCTCATGAGCGCGCATGATGCGACCGTGGTCAGAT 601  
QY 804 CGGAACCTTCTGAGCAATGACCTGTGTCCGAGATCAAGATGCAAGAAAGAGTAAAT 863  
DB 602 TGGAACTTCTGAGCAATGACCTGTGTCCGAGATCAAGATGCAAGAAAGAGTAAAT 661  
QY 864 GGGCTTAAACCTTCCAGTGTTCACCCGAGAAATGTCCTCGGTTTAAAGATGCAAGCG 923  
DB 662 GGGCTTAAACCTTCCAGTGTTCACCCGAGAAATGTCCTCGGTTTAAAGATGCAAGCG 721  
QY 924 CTTCATCTATATAAAGCTGTGTGATCATGAGTGTGTGAGAGGTGAGAGCTCAGCAAC 983  
DB 722 CTTCATCTATATAAAGCTGTGTGATCATGAGTGTGTGAGAGGTGAGAGCTCAGCAAC 781  
QY 984 CCTGATGTCTGCCAATCACTCCAGAGAGCTTCCCTGAGATGAGTCACTAGCTGSCAGTT 1043  
DB 782 CCTGATGTCTGCCAATCACTCCAGAGAGCTTCCCTGAGATGAGTCACTAGCTGSCAGTT 841  
QY 1044 TGTGCTTCTGCAACCTGCGGGGCGAGCGTCTCTCCCACTTCAACCTCTCCCAACG 1103  
DB 842 TGTGCTTCTGCAACCTGCGGGGCGAGCGTCTCTCTCCCACTTCAACCTCTCCCAACG 901  
QY 1104 TGAAGAGAGAGAGAGAGAGGTTGATATCAATCCCGGCTCCAGCAGCAACCCGAGGT 1163  
DB 902 TGAAGAGAGAGAGAGAGGTTGATATCAATCCCGGCTCCAGCAGCAACCCGAGGT 961  
QY 1164 GTTCAAGCTGAGAGACAGAGAGCTTGGAAATGATGCGGGAACTTCAACCTCTCTGCA 1223  
DB 962 GTTCAAGCTGAGAGACAGAGAGCTTGGAAATGATGCGGGAACTTCAACCTCTCTGCA 1021  
QY 1224 AAGCGTGTGACCAAGATGCCCAAGTCCAGGGATCCCTCCGCTGACATTCAGATTTTGT 1283  
DB 1022 AAGCGTGTGACCAAGATGCCCAAGTCCAGGGATCCCTCCGCTGACATTCAGATTTTGT 1081  
QY 1284 CCAACATCCACAAATGAAAGCAATATAATCTACGTGTGACTTGAATGATGAGCAGC 1343  
DB 1082 CCAACATCCACAAATGAAAGCAATATAATCTACGTGTGACTTGAATGATGAGCAGC 1141  
QY 1344 CATGTCACTCACCATGAGCAGCGGCGCTCAAAAGAGCCGCAAGTTGTTCTGCTG 1403  
DB 1142 CATGTCACTCACCATGAGCAGCGGCGCTCAAAAGAGCCGCAAGTTGTTCTGCTG 1201  
QY 1404 TTTTCGTGTTCTTAAAGTCTCGAGCTGAGTGAAGCACTCAACCTCAATCTGCTGCA 1463  
DB 1202 TTTTCGTGTTCTTAAAGTCTCGAGCTGAGTGAAGCACTCAACCTCAATCTGCTGCA 1261  
QY 1464 ACACAAATCTCTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1523  
DB 1262 ACACAAATCTCTCTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1321  
QY 1524 CATTAAGCTGACAGACCGGTAATGCAAGAAATCTTACTCATCTCAGGTGCCAG 1583

DB 1322 CATTAAGCTGACAGACCGGTAATGCAAGAAATCTTACTCATCTCAGGTGCCAG 1381  
QY 1584 TGAACATCTCCACCTGAGCTGAGAGCTGATCACTTCTCTGGAAGCTGAGTCCCA 1643  
DB 1382 TGAACATCTCCACCTGAGCTGAGAGCTGATCACTTCTCTGGAAGCTGAGTCCCA 1441  
QY 1644 GGACAGGCTCAGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1703  
DB 1442 GGACAGGCTCAGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1501  
QY 1704 GGCCTGCAACACCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCA 1763  
DB 1502 GGCCTGCAACACCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCA 1561  
QY 1764 CTTGCGCTCTCTCTGCTCCGAGAGGCTTATCAAGAGATCCAGTGAAGCAGAACTCTC 1823  
DB 1562 CTTGCGCTCTCTCTGCTCCGAGAGGCTTATCAAGAGATCCAGTGAAGCAGAACTCTC 1621  
QY 1824 GGTGACCTTGTGACCTTGTGCTCCAGCTTCCAAAGAGGCTTCAAGCAGGCTGTGAC 1883  
DB 1622 GGTGACCTTGTGACCTTGTGCTCCAGCTTCCAAAGAGGCTTCAAGCAGGCTGTGAC 1681  
QY 1884 GGTGCTCTTATACCTTATTTTAAAGAGAGAGGCTTTTCAAGTGAAGTGTGACCAAA 1943  
DB 1682 GGTGCTCTTATACCTTATTTTAAAGAGAGAGGCTTTTCAAGTGAAGTGTGACCAAA 1741  
QY 1944 AAGCAAGCTCTACCTGAGGACCCCACTGAGGACCGGAGCTTCCATCCCTCAGCTGT 2003  
DB 1742 AAGCAAGCTCTACCTGAGGACCCCACTGAGGACCGGAGCTTCCATCCCTCAGCTGT 1801  
QY 2004 GTCTGGAACATCACTGAGCTGAGGACCGGAGCTTCCATCCCTCAGCTGT 2063  
DB 1802 GTCTGGAACATCACTGAGCTGAGGACCGGAGCTTCCATCCCTCAGCTGT 1861  
QY 2064 GAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2123  
DB 1862 GAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1921  
QY 2124 GAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2183  
DB 1922 GAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1981  
QY 2184 CAGCTTCTGAGTCACTTCTTAACTGAGCCCAAGAGGCTTCAAGCTGAGCTGCT 2243  
DB 1982 CAGCTTCTGAGTCACTTCTTAACTGAGCCCAAGAGGCTTCAAGCTGAGCTGCT 2041  
QY 2244 CTTTCTGAGTCACTTCTTAACTGAGCCCAAGAGGCTTCAAGCTGAGCTGCT 2303  
DB 2042 CTTTCTGAGTCACTTCTTAACTGAGCCCAAGAGGCTTCAAGCTGAGCTGCT 2101  
QY 2304 GGGAGGTGAGTCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2360  
DB 2102 GGGAGGTGAGTCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2158

RESULT 2  
US-09-328-111-370/C  
; Sequence 370 Application US/09328111  
; Patent No. 6262333  
; GENERAL INFORMATION:  
; APPLICANT: Endege, Wilson O.  
; APPLICANT: Steinmann, Kathleen E.  
; APPLICANT: Astle, Jon H.  
; APPLICANT: Burgess, Christopher C.  
; APPLICANT: Bushnell, Steven E.  
; APPLICANT: Carroli III, Eddie  
; APPLICANT: Dertl, Adnan  
; APPLICANT: Ford, Donna M.  
; APPLICANT: Lewis, Marcia E.  
; APPLICANT: Nonahan, John E.  
; APPLICANT: Schlegel, Robert

;; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION  
;; FILE REFERENCE: PRODUCTS  
;; CURRENT APPLICATION NUMBER: US/09/328,111  
;; EARLIER APPLICATION NUMBER: US 60/088,801  
;; NUMBER OF SEQ ID NOS: 850  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 370  
;; LENGTH: 636  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: misc.feature  
;; LOCATION: (1)...(636)  
;; OTHER INFORMATION: n = A,T,C or G  
US-09-328-111-370

Query Match 3.8%; Score 236; DB 3; Length 636;  
Best Local Similarity 99.7%; Pred. No. 3.6e-93;  
Matches 286; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5617 AAGCTTACTGCTGCGCAGCCCTGCTCATCCAGTGTGCTCCATGCTACGAGAAAG 5676  
DB 287 AAGCTTACTGCTGCGCAGCCCTGCTCATCCAGTGTGCTCCATGCTACGAGAAAG 228  
QY 5677 CCAGGCGATGCGAGCTGCTCTAATGCTGTCATGCGACAGAGGAAAGCTTCA 5736  
DB 227 CCAGGCGATGCGAGCTGCTCTAATGCTGTCATGCGACAGAGGAAAGCTTCA 168  
QY 5737 AGGAAGATCACTGCGACAGCAAGCCCAAGCCGACATGCTTGTGAAGTTAGCA 5796  
DB 167 AGGAAGATCACTGCGACAGCAAGCCCAAGCCGACATGCTTGTGAAGTTAGCA 108  
QY 5797 GACTGTGTGTGATGATCTGCAAGTCTTCACTGGAATATTATTATTCATGAGATCTT 5856  
DB 107 GACTGTGTGTGATGATCTGCAAGTCTTCACTGGAATATTATTATTCATGAGATCTT 48  
QY 5857 TTTAGGTGGCATTTTATTCATTCTCCGTGCTTTAATAAACAATGT 5903  
DB 47 TTTAGGTGGCATTTTATTCATTCTCCGTGCTTTAATAAACAATGT 1

## RESULT 3

US-09-313-294A-7155  
; Sequence 7155, Application US/09313294A  
; Patent No. 6476212  
; GENERAL INFORMATION:  
; APPLICANT: Laligudi, Raghunath V.  
; APPLICANT: Ito, Laura Y.  
; APPLICANT: Sherman, Bradley K.  
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR  
; FILE REFERENCE: PL-0017 US  
; CURRENT APPLICATION NUMBER: US/09/313,294A  
; CURRENT FILING DATE: 1999-05-14  
; NUMBER OF SEQ ID NOS: 7600  
; SOFTWARE: PERL Program  
; SEQ ID NO 7155  
; LENGTH: 174  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: misc.feature  
; OTHER INFORMATION: Incyte ID No. 6476212 700381216H1  
US-09-313-294A-7155

Query Match 2.8%; Score 174; DB 4; Length 174;  
Best Local Similarity 100.0%; Pred. No. 4e-66;  
Matches 174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1515 GGAACAAACATTAAGCTGCACAGACACCGGTAAGTCCAAAGAAATCTACTCACTCCA 1574  
|||||

DB 1 GGAACAAACATTAAGCTGCACAGACACCGGTAAGTCCAAAGAAATCTACTCACTCCA 60  
QY 1575 GGTGCCAGTGCATCTCTCAACTGCTGAGAGCTGATGATCTTCTCTGGAAGTGT 1634  
DB 61 GGTGCCAGTGCATCTCTCAACTGCTGAGAGCTGATGATCTTCTCTGGAAGTGT 120  
QY 1635 GGTGCCAAGGACAGCTCAGCTGTGTGCTGTGCGACCCGAGAGCTGCAGCA 1688  
DB 121 GGTGCCAAGGACAGCTCAGCTGTGTGCTGTGCGACCCGAGAGCTGCAGCA 174

## RESULT 4

US-09-292-542A-1/c  
; Sequence 1, Application US/09292542A  
; Patent No. 6531279  
; GENERAL INFORMATION:  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Chumakov, Ilya  
; APPLICANT: Bougueteloret, Lydie  
; TITLE OF INVENTION: Genomic Sequence Of The 5-Lipoxygenase-Activating Protein (FLAP),  
; Patent No. 6531279  
; TITLE OF INVENTION: Polymorphic Markers Thereof And Methods For Detection Of Asthma.  
; FILE REFERENCE: GENSET.026A  
; CURRENT FILING DATE: 1999-04-15  
; CURRENT APPLICATION NUMBER: US/09/292,542A  
; PRIOR APPLICATION NUMBER: US 60/081893  
; PRIOR FILING DATE: 1998-04-15  
; PRIOR APPLICATION NUMBER: US 60/091314  
; PRIOR FILING DATE: 1998-06-30  
; PRIOR APPLICATION NUMBER: US 60/123406  
; PRIOR FILING DATE: 1999-03-08  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: Patent.pm  
; SEQ ID NO 1  
; LENGTH: 43069  
; TYPE: DNA  
; ORGANISM: homo sapiens  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: 1..7708  
; OTHER INFORMATION: potential 5' regulatory region  
; NAME/KEY: misc.feature  
; LOCATION: 36604..43069  
; OTHER INFORMATION: potential 3' regulatory region  
; NAME/KEY: exon  
; LOCATION: 7709..7852  
; OTHER INFORMATION: exon1  
; NAME/KEY: exon  
; LOCATION: 16236..16335  
; OTHER INFORMATION: exon2  
; NAME/KEY: exon  
; LOCATION: 24227..24297  
; OTHER INFORMATION: exon3  
; NAME/KEY: exon  
; LOCATION: 28133..28214  
; OTHER INFORMATION: exon4  
; NAME/KEY: exon  
; LOCATION: 36128..36605  
; OTHER INFORMATION: exon5  
; NAME/KEY: misc.feature  
; LOCATION: 7783..7785  
; OTHER INFORMATION: ATG  
; NAME/KEY: misc.feature  
; LOCATION: 36286..36290  
; OTHER INFORMATION: stop : TAA  
; NAME/KEY: POLYA signal  
; LOCATION: 36581..36586  
; OTHER INFORMATION: AATAAA  
; NAME/KEY: misc.feature  
; LOCATION: 7008..8116  
; OTHER INFORMATION: homology with sequence in ref genbank : M60470  
; NAME/KEY: misc.feature  
; LOCATION: 15995..16549

OTHER INFORMATION: homology with sequence in ref genbank : M63259  
NAME/KEY: misc\_feature  
LOCATION: 24059..24597  
OTHER INFORMATION: homology with sequence in ref genbank : M63260  
NAME/KEY: misc\_feature  
LOCATION: 27875..28412  
OTHER INFORMATION: homology with sequence in ref genbank : M63261  
NAME/KEY: misc\_feature  
LOCATION: 35977..36926  
OTHER INFORMATION: homology with sequence in ref genbank : M63262  
NAME/KEY: misc\_feature  
LOCATION: 7613\_  
OTHER INFORMATION: diverging nucleotide deletion of a A in ref : M60470  
NAME/KEY: misc\_feature  
LOCATION: 16347  
OTHER INFORMATION: diverging nucleotide G in ref : M63259  
NAME/KEY: misc\_feature  
LOCATION: 16348  
OTHER INFORMATION: diverging nucleotide A in ref : M63259  
NAME/KEY: misc\_feature  
LOCATION: 24060  
OTHER INFORMATION: diverging nucleotide deletion of a G in ref : M63260  
NAME/KEY: misc\_feature  
LOCATION: 24067  
OTHER INFORMATION: diverging nucleotide deletion of a G in ref : M63260  
NAME/KEY: misc\_feature  
LOCATION: 27903  
OTHER INFORMATION: diverging nucleotide deletion of a C in ref : M63261  
NAME/KEY: misc\_feature  
LOCATION: 28327  
OTHER INFORMATION: diverging nucleotide deletion of a G in ref : M63261  
NAME/KEY: misc\_feature  
LOCATION: 3851..4189  
OTHER INFORMATION: 10-517  
NAME/KEY: misc\_feature  
LOCATION: 4120..4390  
OTHER INFORMATION: 10-518  
NAME/KEY: misc\_feature  
LOCATION: 4373..4792  
OTHER INFORMATION: 10-253  
NAME/KEY: misc\_feature  
LOCATION: 4814..5043  
OTHER INFORMATION: 10-499  
NAME/KEY: misc\_feature  
LOCATION: 4956..5422  
OTHER INFORMATION: 10-500  
NAME/KEY: misc\_feature  
LOCATION: 5524..5996  
OTHER INFORMATION: 10-522  
NAME/KEY: misc\_feature  
LOCATION: 6218..6672  
OTHER INFORMATION: 10-503  
NAME/KEY: misc\_feature  
LOCATION: 6522..6790  
OTHER INFORMATION: 10-504  
NAME/KEY: misc\_feature  
LOCATION: 7120..7574  
OTHER INFORMATION: 10-204  
NAME/KEY: misc\_feature  
LOCATION: 7513..7933  
OTHER INFORMATION: 10-32  
NAME/KEY: misc\_feature  
LOCATION: 16114..16533  
OTHER INFORMATION: 10-33  
NAME/KEY: misc\_feature  
LOCATION: 24072..24425  
OTHER INFORMATION: 10-34  
NAME/KEY: misc\_feature  
LOCATION: 27978..28401  
OTHER INFORMATION: 10-35  
NAME/KEY: misc\_feature  
LOCATION: 36020..36465  
OTHER INFORMATION: 10-36

NAME/KEY: misc\_feature  
LOCATION: 36318..36659  
OTHER INFORMATION: 10-498  
NAME/KEY: misc\_feature  
LOCATION: 38441..38840  
OTHER INFORMATION: 12-629  
NAME/KEY: misc\_feature  
LOCATION: 42235..42749  
OTHER INFORMATION: 12-628 complement  
NAME/KEY: allele  
LOCATION: 3950  
OTHER INFORMATION: 10-517-100 : polymorphic base S  
NAME/KEY: allele  
LOCATION: 4243  
OTHER INFORMATION: 10-518-125 : polymorphic base K  
NAME/KEY: allele  
LOCATION: 4312  
OTHER INFORMATION: 10-518-194 : polymorphic base R  
NAME/KEY: allele  
LOCATION: 4490  
OTHER INFORMATION: 10-253-118 : polymorphic base R  
NAME/KEY: allele  
LOCATION: 4670  
OTHER INFORMATION: 10-253-298 : polymorphic base S  
NAME/KEY: allele  
LOCATION: 4687  
OTHER INFORMATION: 10-253-315 : polymorphic base Y  
NAME/KEY: allele  
LOCATION: 4968  
OTHER INFORMATION: 10-499-155 : polymorphic base R  
NAME/KEY: allele  
LOCATION: 5140  
OTHER INFORMATION: 10-500-185 : polymorphic base Y  
NAME/KEY: allele  
LOCATION: 5213  
OTHER INFORMATION: 10-500-258 : polymorphic base K  
NAME/KEY: allele  
LOCATION: 5364  
OTHER INFORMATION: 10-500-410 : polymorphic base R  
NAME/KEY: allele  
LOCATION: 5594  
OTHER INFORMATION: 10-522-71 : polymorphic base R  
NAME/KEY: allele  
LOCATION: 6370  
OTHER INFORMATION: 10-503-159 : polymorphic base K  
NAME/KEY: allele  
LOCATION: 6693  
OTHER INFORMATION: 10-504-172 : polymorphic base W  
NAME/KEY: allele  
LOCATION: 6763  
OTHER INFORMATION: 10-504-243 : polymorphic base M  
NAME/KEY: allele  
LOCATION: 7445  
OTHER INFORMATION: 10-204-326 : polymorphic base R  
NAME/KEY: allele  
LOCATION: 7870  
OTHER INFORMATION: 10-32-357 : polymorphic base M  
NAME/KEY: allele  
LOCATION: 16288  
OTHER INFORMATION: 10-33-175 : polymorphic base Y  
NAME/KEY: allele  
LOCATION: 16347  
OTHER INFORMATION: 10-33-234 : polymorphic base M  
NAME/KEY: allele  
LOCATION: 16383  
OTHER INFORMATION: 10-33-270 : polymorphic base R  
NAME/KEY: allele

Query Match 0.9% ; Score 53 ; DB 4 ; Length 43069 ;  
Best Local Similarity 100.0% ; Pred. No. 1.8e-13 ;  
Matches 53 ; Conservative 0 ; Mismatches 0 ; Indels 0 ; Gaps 0 ;

Qy 4547 AGCTGAGTGCAGTGGCGCAATCTGGCTACATGCAACTTGTGCTCTGCG 4599

Db 28976 AGCGTGAGTGCAGCGCCCAATCTCGCTCACTGCAACCTTGCTCTGGG 28924

US-09-146-053-6/c  
Sequence 6, Application US/09146053A  
Patent No. 6399349

GENERAL INFORMATION:

APPLICANT: Ryan, James W.

APPLICANT: Sprinkle, Terry Joe Curtis

APPLICANT: Venema, Richard C.

TITLE OF INVENTION: Human Aminopeptidase P Gene

FILE REFERENCE: MCG103

CURRENT APPLICATION NUMBER: US/09/146,053A

CURRENT FILING DATE: 1998-09-02

EARLIER APPLICATION NUMBER: 60/057,854

EARLIER FILING DATE: 1997-09-02

NUMBER OF SEQ ID NOS: 7

SOFTWARE: Patent In Ver. 2.0

SEQ ID NO 6

LENGTH: 45546

TYPE: DNA

ORGANISM: Homo sapiens

US-09-146-053-6

Query Match 0.8%; Score 51; DB 4; Length 45546;  
Best Local Similarity 100.0%; Pred. No. 1.3e-12;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4671 TTTTGTATTTTGTAGAGACAGAGGTTTCACCATGTTGCCAGGCTGTCT 4721  
DB 39160 TTTTGTATTTTGTAGAGACAGAGGTTTCACCATGTTGCCAGGCTGTCT 39110

US-09-851-896-3  
Sequence 3, Application US/09851896  
Patent No. 6410325

GENERAL INFORMATION:

APPLICANT: C. Frank Bennett

APPLICANT: Susan M. Freiler

APPLICANT: Andrew T. Watt

TITLE OF INVENTION: ANTISENSE MODULATION OF PHOSPHOLIPASE A2, GROUP VI (CA2+-INDEPENDENT

TITLE OF INVENTION: EXPRESSION

FILE REFERENCE: RTS-0220

CURRENT APPLICATION NUMBER: US/09/851,896

CURRENT FILING DATE: 2001-05-08

NUMBER OF SEQ ID NOS: 89

SEQ ID NO 3

LENGTH: 70000

TYPE: DNA

ORGANISM: Homo sapiens

US-09-851-896-3

Query Match 0.8%; Score 49; DB 4; Length 70000;  
Best Local Similarity 100.0%; Pred. No. 9.5e-12;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4547 AGCGTGAGTGCAGCGCCCAATCTCGCTCACTGCAACCTTGCTCTCC 4595  
DB 11896 AGCGTGAGTGCAGCGCCCAATCTCGCTCACTGCAACCTTGCTCTCC 11944

US-08-994-076-1/c  
Sequence 1, Application US/08994076  
Patent No. 6500937

GENERAL INFORMATION:

APPLICANT: Ervin, Jr., Paul R.

TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCE

TITLE OF INVENTION: OF MAMMASTATIN AND METHODS OF USE

NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt  
STREET: 3100 No. 6500937 West Center, 90 South 7th Street  
CITY: Minneapolis  
STATE: MN

COUNTRY: USA

ZIP: 55402

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/994,076

FILING DATE: 19-DEC-1997

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/027,315

FILING DATE: 03-OCT-1996

APPLICATION NUMBER: PCT/US97/18026

FILING DATE: 03-OCT-1997

APPLICATION NUMBER: 08/943,828

FILING DATE: 03-OCT-1997

ATTORNEY/AGENT INFORMATION:

NAME: Ketteiberg, Denise M

REGISTRATION NUMBER: 33,924

REFERENCE/DOCKET NUMBER: 4273.1US11

TELECOMMUNICATION INFORMATION:

TELEPHONE: 612-332-5100

TELEFAX: 612-332-9081

TELEX:

SEQUENCE CHARACTERISTICS:

LENGTH: 2326 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: Genomic DNA

FEATURE:

NAME/KEY: Coding Sequence

LOCATION: 348...1961

OTHER INFORMATION:

US-08-994-076-1

Query Match 0.7%; Score 45; DB 4; Length 2326;  
Best Local Similarity 100.0%; Pred. No. 6e-10;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4678 TTTTGTAGAGACAGAGGTTTCACCATGTTGCCAGGCTGTCTC 4722  
DB 2144 TTTTGTAGAGACAGAGGTTTCACCATGTTGCCAGGCTGTCTC 2100

US-09-285-379-1/c  
Sequence 1, Application US/09285379  
Patent No. 6451765

GENERAL INFORMATION:

APPLICANT: Ervin Jr., Paul R.

TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCE OF MAMMASTATIN AND

TITLE OF INVENTION: METHODS OF USE

FILE REFERENCE: 4273.1US11

CURRENT APPLICATION NUMBER: US/09/285,379

CURRENT FILING DATE: 1999-04-02

EARLIER APPLICATION NUMBER: 08/943,828

EARLIER FILING DATE: 1997-10-03

EARLIER APPLICATION NUMBER: 60/027,315

EARLIER FILING DATE: 1996-10-03

NUMBER OF SEQ ID NOS: 7

SOFTWARE: Patent In Ver. 2.0

SEQ ID NO 1

LENGTH: 2418

TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-285-379-1

Query Match  
Best Local Similarity 100.0%; Score 45; DB 4; Length 2418;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4678 TTTTGTAGAGACAGAGGTTTCCATGTTGCCAGGCTGCTC 4722  
DB 2233 TTTTGTAGAGACAGAGGTTTCCATGTTGCCAGGCTGCTC 2189

RESULT 9  
US-09-369-912-1/c  
Sequence 1, Application US/09369912  
Patent No. 6492504  
GENERAL INFORMATION:  
APPLICANT: Evin, Jr., Paul R.  
TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCE OF MAMMASTATIN AND  
FILE REFERENCE: 4273.1USU1  
CURRENT APPLICATION NUMBER: US/09/369,912  
EARLIER FILING DATE: 1999-08-06  
EARLIER APPLICATION NUMBER: 60/027,315  
NUMBER OF SEQ ID NOS: 7  
SOFTWARE: Patentin Ver. 2.0  
SEO ID NO 1  
LENGTH: 2418  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-369-912-1

Query Match  
Best Local Similarity 100.0%; Score 45; DB 4; Length 2418;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4678 TTTTGTAGAGACAGAGGTTTCCATGTTGCCAGGCTGCTC 4722  
DB 2233 TTTTGTAGAGACAGAGGTTTCCATGTTGCCAGGCTGCTC 2189

RESULT 10  
US-08-975-080-35/c  
Sequence 35, Application US/08975080  
Patent No. 6245523  
GENERAL INFORMATION:  
APPLICANT: Allieri, Dario C.  
TITLE OF INVENTION: SURVIVIN, A PROTEIN THAT INHIBITS  
CELLULAR APOPTOSIS, AND ITS MODULATION  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP  
STREET: 1800 M Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20036-5869  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/975,080  
FILING DATE: 20-NOV-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/031,435  
FILING DATE: 20-NOV-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Adler, Reid G.  
REGISTRATION NUMBER: 30,988

REFERENCE/DOCKET NUMBER: 044574-5022-01-WO  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-467-7000  
TELEFAX: 202-467-7176  
INFORMATION FOR SEQ ID NO: 35:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14796 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-975-080-35

Query Match  
Best Local Similarity 100.0%; Score 45; DB 3; Length 14796;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4747 CCTGCTGGGCTCCCAAGTGTGGATTACAGATGAGCCAC 4791  
DB 1410 CCTGCTGGGCTCCCAAGTGTGGATTACAGATGAGCCAC 1366

RESULT 11  
US-09-630-706-10/c  
Sequence 10, Application US/09630706  
Patent No. 6277640  
GENERAL INFORMATION:  
APPLICANT: C. Frank Bennett  
TITLE OF INVENTION: ANTISENSE MODULATION OF HER-3 EXPRESSION  
FILE REFERENCE: RTS-0053  
CURRENT APPLICATION NUMBER: US/09/630,706  
CURRENT FILING DATE: 2000-08-01  
NUMBER OF SEQ ID NOS: 94  
SEO ID NO 10  
LENGTH: 14796  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (2811)...(2921)  
NAME/KEY: CDS  
LOCATION: (3174)...(3283)  
NAME/KEY: CDS  
LOCATION: (5158)...(5275)  
NAME/KEY: CDS  
LOCATION: (11955)...(12044)  
US-09-630-706-10

Query Match  
Best Local Similarity 100.0%; Score 45; DB 3; Length 14796;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4747 CCTGCTGGGCTCCCAAGTGTGGATTACAGATGAGCCAC 4791  
DB 1410 CCTGCTGGGCTCCCAAGTGTGGATTACAGATGAGCCAC 1366

RESULT 12  
US-09-496-694B-3/c  
Sequence 3, Application US/09496694B  
Patent No. 6335194  
GENERAL INFORMATION:  
APPLICANT: C. Frank Bennett  
APPLICANT: Elizabeth J. Ackermann  
APPLICANT: Eric E. Swayze  
APPLICANT: Lex M. Cowsett  
TITLE OF INVENTION: ANTISENSE MODULATION OF SURVIVIN EXPRESSION  
FILE REFERENCE: ISPH-0439  
CURRENT APPLICATION NUMBER: US/09/496,694B  
CURRENT FILING DATE: 2000-02-02  
PRIOR APPLICATION NUMBER: 09/286,407  
PRIOR FILING DATE: 1999-04-05

```
/ PRIOR APPLICATION NUMBER: 09/163,162
/ PRIOR FILING DATE: 1998-09-29
/ NUMBER OF SEQ ID NOS: 249
/ SEQ ID NO 3
/ LENGTH: 14796
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (2811)...(2921)
/ NAME/KEY: CDS
/ LOCATION: (3174)...(3283)
/ NAME/KEY: CDS
/ LOCATION: (5158)...(5275)
/ NAME/KEY: CDS
/ LOCATION: (11955)...(12044)
US-09-496-694B-3

Query Match
Best Local Similarity 100.0%; Score 45; DB 4; Length 14796;
Pred. No. 5.6e-10;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

4747 CCTGCTCGGCTCCCAAGTGTGGATTACAGATGTAGCCAC 4791
|||||
1410 CCTGCTCGGCTCCCAAGTGTGGATTACAGATGTAGCCAC 1366

RESULT 13
US-09-735-934A-3/C
/ Sequence 3, Application US/09735934A
/ Patent No. 6372468
/ GENERAL INFORMATION:
/ APPLICANT: Li, JiaYin et al
/ TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
/ TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
/ FILE REFERENCE: CL000851
/ CURRENT APPLICATION NUMBER: US/09/735,934A
/ CURRENT FILING DATE: 2000-12-14
/ NUMBER OF SEQ ID NOS: 4
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 3
/ LENGTH: 43950
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-735-934A-3

Query Match
Best Local Similarity 100.0%; Score 44; DB 4; Length 43950;
Pred. No. 1.5e-09;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

4678 TTTTGTAGAGACAGGGTTTCAACCATGTGGCCAGGCTGTCT 4721
|||||
14702 TTTTGTAGAGACAGGGTTTCAACCATGTGGCCAGGCTGTCT 14659

RESULT 14
US-10-060-332-3/C
/ Sequence 3, Application US/10060332
/ Patent No. 6528294
/ GENERAL INFORMATION:
/ APPLICANT: Li, JiaYin et al
/ TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
/ TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
/ FILE REFERENCE: CL000851DIV
/ CURRENT APPLICATION NUMBER: US/10/060,332
/ CURRENT FILING DATE: 2002-02-01
/ NUMBER OF SEQ ID NOS: 4
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 3
/ LENGTH: 43950
/ TYPE: DNA
```

```
/ ORGANISM: Homo sapiens
US-10-060-332-3

Query Match
Best Local Similarity 100.0%; Score 44; DB 4; Length 43950;
Pred. No. 1.5e-09;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

4678 TTTTGTAGAGACAGGGTTTCAACCATGTGGCCAGGCTGTCT 4721
|||||
14702 TTTTGTAGAGACAGGGTTTCAACCATGTGGCCAGGCTGTCT 14659

RESULT 15
US-09-128-155-17/C
/ Sequence 17, Application US/09128155
/ Patent No. 6117654
/ GENERAL INFORMATION:
/ APPLICANT: Pan, Yang
/ TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY
/ TITLE OF INVENTION: AND USES THEREOF
/ FILE REFERENCE: 09404/052001
/ CURRENT APPLICATION NUMBER: US/09/128,155
/ CURRENT FILING DATE: 1998-08-03
/ EARLIER APPLICATION NUMBER: US 60/091,650
/ EARLIER FILING DATE: 1998-07-02
/ EARLIER APPLICATION NUMBER: US 60/054,646
/ EARLIER FILING DATE: 1997-08-04
/ NUMBER OF SEQ ID NOS: 18
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 17
/ LENGTH: 176373
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(176373)
/ OTHER INFORMATION: n = A,T,C or G
US-09-128-155-17

Query Match
Best Local Similarity 100.0%; Score 44; DB 3; Length 176373;
Pred. No. 1.4e-09;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

4750 GCCTGGGCTCCCAAGTGTGGATTACAGATGTAGCCACCG 4793
|||||
168946 GCCTGGGCTCCCAAGTGTGGATTACAGATGTAGCCACCG 168903

Search completed: February 20, 2004, 13:55:05
Job time : 340 secs
```

**This Page Blank (uspto)**





```

; LOCATION: (2794) .. (6163)
; NAME/KEY: CDS
; LOCATION: (283) .. (2793)
US-09-899-569A-3

```

Query Match	Score	DB	Length
100.0%	6163	10	6163

Matches 6163; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	CCAAACGCGCAATGGGGAGTAGTAGGGAACCCAGCAACCCGGTGCGGAGGCTTGCAACC	60
Db	1	CCAAACGCGCAATGGGGAGTAGTAGGGAACCCAGCAACCCGGTGCGGAGGCTTGCAACC	60
QY	61	TGGAGAGGAGAGGCGGTGCTGTAGAGCGAGAAAGAGAGAGAGAGAGAGAGAGAGCGAC	120
Db	61	TGGAGAGGAGAGAGCGGTGCTGTAGAGCGAGAAAGAGAGAGAGAGAGAGAGAGAGCGAC	120
QY	121	CGGGTCAGCTCGGAGATCTGTGTGCGAGAGGCGGGCTGGAGCCGGTCCGCCCGCGAG	180
Db	121	CGGGTCAGCTCGGAGATCTGTGTGCGAGAGGCGGGCTGGAGCCGGTCCGCCCGCGAG	180
QY	181	GTGAGTGAGCAGGGGGGAGCGGAGCTGTGCGCGGGCTTGGGGCTCTGGGGCGCGCGCC	240
Db	181	GTGAGTGAGCAGGGGGGAGCGGAGCTGTGCGCGGGCTTGGGGCTCTGGGGCGCGCGCC	240
QY	241	CCACCGTCGTTTTCCCAACCGAGGCGGAGCGCTCCCGAGTCAATGGCGGCGCTGAACCTGC	300
Db	241	CCACCGTCGTTTTCCCAACCGAGGCGGAGCGCTCCCGAGTCAATGGCGGCGCTGAACCTGC	300
QY	301	GGGGTCTTATCGCACTGTGTAAGGGTTCTGTCTGTGGGTGCGGCGCGCTGCGCGGG	360
Db	301	GGGGTCTTATCGCACTGTGTAAGGGTTCTGTCTGTGGGTGCGGCGCGCTGCGCGGG	360
QY	361	GCAGAACTTTGAGATTGTCTCTGCCAGAGAAACCAATTACAGTTCTATAAAGCTG	420
Db	361	GCAGAACTTTGAGATTGTCTCTGCCAGAGAAACCAATTACAGTTCTATAAAGCTG	420
QY	421	GGGAGCCCGACTGTGCTGGCAAAACCCTGTATCATGTGCACTTTCAAAAGCATATAACC	480
Db	421	GGGAGCCCGACTGTGCTGGCAAAACCCTGTATCATGTGCACTTTCAAAAGCATATAACC	480
QY	481	ATGTTGTCCATCAAGCTCTGAGAAAGAAATAGCTTTTACCTTTAGCTGCAGAGTCTGAG	540
Db	481	ATGTTGTCCATCAAGCTCTGAGAAAGAAATAGCTTTTACCTTTAGCTGCAGAGTCTGAG	540
QY	541	AATCATCTTGTCTATAGAGATCCAGAAAAATATTGACTGATGTCAAGGCCATGTCTTTT	600
Db	541	AATCATCTTGTCTATAGAGATCCAGAAAAATATTGACTGATGTCAAGGCCATGTCTTTT	600
QY	601	GGGAGGTTCAAGTTAGAGCATCGGTTTGTGCTTACCCTCAACAGAACTTTATC	660
Db	601	GGGAGGTTCAAGTTAGAGCATCGGTTTGTGCTTACCCTCAACAGAACTTTATC	660
QY	661	TGGAGTGTCAAAAGCTATTAAGAGCATCGGTTTGTGCTTACCCTCAACAGAACTTTATC	720
Db	661	TGGAGTGTCAAAAGCTATTAAGAGCATCGGTTTGTGCTTACCCTCAACAGAACTTTATC	720
QY	721	AGGCAGATCGTCTCGGGTGAAGAGCTGCCAGACGAGTCACTCACTCCATCAGCGGCGA	780
Db	721	AGGCAGATCGTCTCGGGTGAAGAGCTGCCAGACGAGTCACTCACTCCATCAGCGGCGA	780
QY	781	ATCGATGCCAACCCTGTCAAGATCGGAACCTTCTGACGAATGGCATGTGTCCCGGATC	840
Db	781	ATCGATGCCAACCCTGTCAAGATCGGAACCTTCTGACGAATGGCATGTGTCCCGGATC	840
QY	841	AAATATGCAAGAAAGAGATGAAAATGGCTTTACACTTCCCATGGTTTCAACCCAGAAATGTC	900
Db	841	AAATATGCAAGAAAGAGATGAAAATGGCTTTACACTTCCCATGGTTTCAACCCAGAAATGTC	900
QY	901	TTCGGCTTCAGACTTGCAAAACCGCTCATTTAAAGCGTCTGTGATCATGAGTCTGTG	960
Db	901	TTCGGCTTCAGACTTGCAAAACCGCTCATTTAAAGCGTCTGTGATCATGAGTCTGTG	960

OY	961	TTTAAAGGTGAAGGCTCAGCAACCCTGATGTCTGGCCAACTAACCGAAGAGCTTCCCTGAG	1020
Ds	961	TTTAAAGGTGAAGGCTCAGCAACCCTGATGTCTGGCCAACTAACCGAAGAGCTTCCCTGAG	1020
OY	1021	GATAGGCTCATGAGTGGCAGATTGTGTTCTCTGCAACCTGCGGGGCGAGGCTCTCTTC	1080
Ds	1021	GATAGGCTCATGAGTGGCAGATTGTGTTCTCTGCAACCTGCGGGGCGAGGCTCTCTTC	1080
OY	1081	CTCAACTTCAAACCTCTCCAACTGTGAGAGGAGGAGCGGGTTGAATCTACATCCG	1140
Ds	1081	CTCAACTTCAAACCTCTCCAACTGTGAGAGGAGGAGCGGGTTGAATCTACATCCG	1140
OY	1141	GGCTTCCACCCCAACCCCGAGGTGTTAACTGTGAGAGCAAGCAAGCCTGGGAACATAGGC	1200
Ds	1141	GGCTTCCACCCCAACCCCGAGGTGTTAACTGTGAGAGCAAGCAAGCCTGGGAACATAGGC	1200
OY	1201	GGAAACTTCAAACCTCTCTCTGCAAGGCTGTGACCAAGATGCCCAAGTCCAGGGATCCTC	1260
Ds	1201	GGAAACTTCAAACCTCTCTCTCTGCAAGGCTGTGACCAAGATGCCCAAGTCCAGGGATCCTC	1260
OY	1261	CGGCTGCAGTTCCAAAGTTTGTGTCACATCCACAAATGAAAGCAATAAATCTACGTG	1320
Ds	1261	CGGCTGCAGTTCCAAAGTTTGTGTCACATCCACAAATGAAAGCAATAAATCTACGTG	1320
OY	1321	GTGATCTTGAAGTAATGAGCGAGCCATGTCACTACCATCGAGCCACGGCCCTCAACAG	1380
Ds	1321	GTGATCTTGAAGTAATGAGCGAGCCATGTCACTACCATCGAGCCACGGCCCTCAACAG	1380
OY	1381	AGCCGCAAGTTTGTCCCTGGCTGTTTGTGTGTATCAATCTGGGACCTGCAGTAGCAAC	1440
Ds	1381	AGCCGCAAGTTTGTCCCTGGCTGTTTGTGTGTATCAATCTGGGACCTGCAGTAGCAAC	1440
OY	1441	CTCACCTGACATCTGAGCTCCAAACACAAATCTCTCTTGTGTGATCTTGACAGT	1500
Ds	1441	CTCACCTGACATCTGAGCTCCAAACACAAATCTCTCTTGTGTGATCTTGACAGT	1500
OY	1501	CTGTGATGATGTGGAAAAAACATTAAGCTGCAACCAACCGGTACTGCGCAAGGAAA	1560
Ds	1501	CTGTGATGATGTGGAAAAAACATTAAGCTGCAACCAACCGGTACTGCGCAAGGAAA	1560
OY	1561	TTCATCTCACTCCAGGAGGCCAAGTACATCTCCACCTGGCTGTGGAGCTGCATGACTTC	1620
Ds	1561	TTCATCTCACTCCAGGAGGCCAAGTACATCTCCACCTGGCTGTGGAGCTGCATGACTTC	1620
OY	1621	TCTGGAAGCTGTGTGTGTGCCAAGACAGGCTCAGGCTGTGTGTGTGTGCCAGCCCAAG	1680
Ds	1621	TCTGGAAGCTGTGTGTGTGCCAAGACAGGCTCAGGCTGTGTGTGTGTGCCAGCCCAAG	1680
OY	1681	CTGCAGCAGATACACACGAAAGCCCTGCAACCAAGCTTCAGCTACTCTGTGTGGCAGT	1740
Ds	1681	CTGCAGCAGATACACACGAAAGCCCTGCAACCAAGCTTCAGCTACTCTGTGTGGCAGT	1740
OY	1741	GCCATTAACCGACGAGCCTGTACTTCTGGGCTCCTTGGCCCGGAGGCTCTATCAACAG	1800
Ds	1741	GCCATTAACCGACGAGCCTGTACTTCTGGGCTCCTTGGCCCGGAGGCTCTATCAACAG	1800
OY	1801	ATCAGGTGAAGCAGAACATCTCGGTGACCTTGCGCACTTTGGCCCCCAGCTTCCACAA	1860
Ds	1801	ATCAGGTGAAGCAGAACATCTCGGTGACCTTGCGCACTTTGGCCCCCAGCTTCCACAA	1860
OY	1861	GAGGCTTCCAGGCAAGGCTTGAACGGTGTCTTTATATCTTAATTTCAAGAGAAAGCGTT	1920
Ds	1861	GAGGCTTCCAGGCAAGGCTTGAACGGTGTCTTTATATCTTAATTTCAAGAGAAAGCGTT	1920
OY	1921	TTTACCGGTGACCCCTGACAAAAAAGCAAGGTCTAACCTGAGAACCCCAACTGGGACCG	1980
Ds	1921	TTTACCGGTGACCCCTGACAAAAAAGCAAGGTCTAACCTGAGAACCCCAACTGGGACCG	1980
OY	1981	GGGCTGTGCATCCCTCACTCTGTGTCTCTGGAATATGAGCGTGGCCACAGACAGAGTGAGC	2040
Ds	1981	GGGCTGTGCATCCCTCACTCTGTGTCTCTGGAATATGAGCGTGGCCACAGACAGAGTGAGC	2040
OY	2041	TGCTGTGATTTCTTTAAGAGCGGAGGAGGTGTGTGCGACAGACAGGCGCGCATTTATG	2100

2041 TGCCGATCTTTCTTAAGAGCGGAGCGCGGTGCTGCGAGACGAGCGGCGATTCAATG 2100  
2101 ATCATTCGAGAGACGCGGACCGGGCTGAGAGATCTTCAGCGTGGAGAGAGATGTGTC 2160  
2101 ATCATTCGAGAGACGCGGACCGGGCTGAGAGATCTTCAGCGTGGAGAGAGATGTGTC 2160  
2161 CCCAAGCCAAAGCTTCACCATCACAGCTTCTGGGTCAACATCTCTAACTGAGCCCCAG 2220  
2161 CCCAAGCCAAAGCTTCACCATCACAGCTTCTGGGTCAACATCTCTAACTGAGCCCCAG 2220  
2221 AGCGGCAAGAGCTAGAGCTGCTCTTCTGGGTGACATTTACCCCAAGGACTGTGACTTG 2280  
2221 AGCGGCAAGAGCTAGAGCTGCTCTTCTGGGTGACATTTACCCCAAGGACTGTGACTTG 2280  
2281 ACTGTCATCTCATGCGAGCGGTGGAGGTGAGTCTTACTGCTGTCTGCGCTGCGGCTC 2340  
2281 ACTGTCATCTCATGCGAGCGGTGGAGGTGAGTCTTACTGCTGTCTGCGCTGCGGCTC 2340  
2341 ATCATTTGCTGTGAAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 2400  
2341 ATCATTTGCTGTGAAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 2400  
2401 TACAATGGCAATCACTAATCTGAGATGCGGAGGAGCCAAAAAGTTTCAAGAAAGGCGA 2460  
2401 TACAATGGCAATCACTAATCTGAGATGCGGAGGAGCCAAAAAGTTTCAAGAAAGGCGA 2460  
2461 AAGGACATGACTCCCATGTGTATGAGTCACTCGAGACACCATGTGTATGTGGGATCTG 2520  
2461 AAGGACATGACTCCCATGTGTATGAGTCACTCGAGACACCATGTGTATGTGGGATCTG 2520  
2521 CTACAGATTCACAGCGGCTCTCTCTGAGCGAGAGTGGACACCTTACCGGCGGTTCCAG 2580  
2521 CTACAGATTCACAGCGGCTCTCTCTGAGCGAGAGTGGACACCTTACCGGCGGTTCCAG 2580  
2581 GGCAACCATGGGGGTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2640  
2581 GGCAACCATGGGGGTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2640  
2641 AAGTTGGGCACTGAGAGAGCCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2700  
2641 AAGTTGGGCACTGAGAGAGCCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2700  
2701 ACCTTCTCCATCCCAATGAGGATGTAAAGCAGACAGACAGACATTCCTTACTG 2760  
2701 ACCTTCTCCATCCCAATGAGGATGTAAAGCAGACAGACAGACATTCCTTACTG 2760  
2761 AACACTCAGAGCCCATGAGGAGCAGACATTAATTGATTCATTCAGACGCTTGTGTA 2820  
2761 AACACTCAGAGCCCATGAGGAGCAGACATTAATTGATTCATTCAGACGCTTGTGTA 2820  
2821 GTTTCATTAAGCAGGCGACTGAGACACCGGTCTGTCTCTCTCTCTCTCTCTCTCTCTCT 2880  
2821 GTTTCATTAAGCAGGCGACTGAGACACCGGTCTGTCTCTCTCTCTCTCTCTCTCTCTCT 2880  
2881 GAGGAATTAATCAGAAAGAACGAGAGGTTTCTCTGAGACACCGGCACTTCACTGCG 2940  
2881 GAGGAATTAATCAGAAAGAACGAGAGGTTTCTCTGAGACACCGGCACTTCACTGCG 2940  
2941 TCAGTGAATCTATTCTAAGGCGAAGCATTTGAAATGATGAAATTCGAATCTGATACAGT 3000  
2941 TCAGTGAATCTATTCTAAGGCGAAGCATTTGAAATGATGAAATTCGAATCTGATACAGT 3000  
3001 CATGACAGCTCATGTGCT 3060  
3001 CATGACAGCTCATGTGCT 3060  
3061 AGAGAGGCTTAGATCTGCT 3120  
3061 AGAGAGGCTTAGATCTGCT 3120  
3121 AGAGGCTTGCCT 3180  
3121 AGAGGCTTGCCT 3180

3121 AGAGGCTTGCCT 3180  
3181 CTCTCACTGGGGGTCCCGAGATGAAAGACGACATGTGCTTTTATTAATTAATTAATTTG 3240  
3181 CTCTCACTGGGGGTCCCGAGATGAAAGACGACATGTGCTTTTATTAATTAATTAATTTG 3240  
3241 TGGTCTGTGTTATTAAGAGATCAAAATGTAAACACATAGCTCTTTTCACTGACTTA 3300  
3241 TGGTCTGTGTTATTAAGAGATCAAAATGTAAACACATAGCTCTTTTCACTGACTTA 3300  
3301 GTAATTAATCTACTACTGTTGATGCTGGGTGTGACTTCTTACTGACCGCTAAT 3360  
3301 GTAATTAATCTACTACTGTTGATGCTGGGTGTGACTTCTTACTGACCGCTAAT 3360  
3361 AAAGGTGCTGTGCTCCCGAGGTGGGATTAATTAATTAATTAATTAATTAATTAATTAAT 3420  
3361 AAAGGTGCTGTGCTCCCGAGGTGGGATTAATTAATTAATTAATTAATTAATTAATTAAT 3420  
3421 ATGTGTGTTTGAAGCAGATTTGACATATCTGCTTTGATTAAGAGACTTCTGATTTCTC 3480  
3421 ATGTGTGTTTGAAGCAGATTTGACATATCTGCTTTGATTAAGAGACTTCTGATTTCTC 3480  
3481 TAGGTGCTGTGTTATTCCTCATTTGGAATTCATTTGAAATCCCATTTGCTATAGT 3540  
3481 TAGGTGCTGTGTTATTCCTCATTTGGAATTCATTTGAAATCCCATTTGCTATAGT 3540  
3541 CCTAGCAATTAAGAAATTTCTCAAGTTTCTCATGTGCGGTCTCTTACGTCGAGCAATA 3600  
3541 CCTAGCAATTAAGAAATTTCTCAAGTTTCTCATGTGCGGTCTCTTACGTCGAGCAATA 3600  
3601 CTTTGACATTTAAAGAGAAATTTAGAGATATTTCTCATCTCTTAAATATTAATTAATTA 3660  
3601 CTTTGACATTTAAAGAGAAATTTAGAGATATTTCTCATCTCTTAAATATTAATTAATTA 3660  
3661 TACCAAAAGTGGCCCCCTGATTAATTTCTGTGCGACATGCAACCATTAATCTGTTAG 3720  
3661 TACCAAAAGTGGCCCCCTGATTAATTTCTGTGCGACATGCAACCATTAATCTGTTAG 3720  
3721 CTTAAAAACAACATTAAGCTTATAGTCTGAGGATCAAGATTCGAAATGAGATGCTCT 3780  
3721 CTTAAAAACAACATTAAGCTTATAGTCTGAGGATCAAGATTCGAAATGAGATGCTCT 3780  
3781 GAATGAATCAAGGTGTCAAGAGCTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3840  
3781 GAATGAATCAAGGTGTCAAGAGCTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3840  
3841 GTTCTTGCATTTCAAGCTTCTAAGGCTGGCTGATTTCCAGGCTCCAGTGGCTGCTC 3900  
3841 GTTCTTGCATTTCAAGCTTCTAAGGCTGGCTGATTTCCAGGCTCCAGTGGCTGCTC 3900  
3901 AAGCTTTTCTCAATGAGATCACTGTGACACTGGCCCTCCCACTTCTCTTTGACTGAC 3960  
3901 AAGCTTTTCTCAATGAGATCACTGTGACACTGGCCCTCCCACTTCTCTTTGACTGAC 3960  
3961 AAAGCCACAGAGAAATCCAGAGTAAATCTCTCATCTTAAAGATCTTCACTATCTGGA 4020  
3961 AAAGCCACAGAGAAATCCAGAGTAAATCTCTCATCTTAAAGATCTTCACTATCTGGA 4020  
4021 AAGGCTTTTGCATGCAAGCAATTAAGCAGAGTGGGATTAAGACAGACATCTT 4080  
4021 AAGGCTTTTGCATGCAAGCAATTAAGCAGAGTGGGATTAAGACAGACATCTT 4080  
4081 TGGGGTGTGTTATTTCTGCTTACCAACATTTCTGCACTTCCACAGAGAGGCTA 4140  
4081 TGGGGTGTGTTATTTCTGCTTACCAACATTTCTGCACTTCCACAGAGAGGCTA 4140  
4141 CAAATGATCTGGGCGACAGGATGTTTGTGTTAGCTTGGGACTCTTAACCTTAAAAA 4200  
4141 CAAATGATCTGGGCGACAGGATGTTTGTGTTAGCTTGGGACTCTTAACCTTAAAAA 4200  
4201 ACCCAATCAGAAAGATCTGGCAATGCTGGGCTCAATTTCTCACTTGAACAACTGGC 4260  
4201 ACCCAATCAGAAAGATCTGGCAATGCTGGGCTCAATTTCTCACTTGAACAACTGGC 4260  
4201 ACCCAATCAGAAAGATCTGGCAATGCTGGGCTCAATTTCTCACTTGAACAACTGGC 4260

4261 TGGAGCTGGGACACAGCTCTGCTTTAGAGGGGTGTCATTACACAGGTCCACAGC 4320  
4261 TGGAGCTGGGACACAGCTCTGCTTTAGAGGGGTGTCATTACACAGGTCCACAGC 4320  
4321 CCACATAGAGGCTTACACTTCCCAATGAGGCTAAGGTTGTTTCTAATGATCAAG 4380  
4321 CCACATAGAGGCTTACACTTCCCAATGAGGCTAAGGTTGTTTCTAATGATCAAG 4380  
4381 CCCCTGAGGTTGATTTATTTGTAATGAAAGAAAGAAAGCTGGATTAATCTAATCAG 4440  
4381 CCCCTGAGGTTGATTTATTTGTAATGAAAGAAAGAAAGCTGGATTAATCTAATCAG 4440  
4441 TGAATGACATGAGACCAATGTGTCTCATTTACATTTCTTTTCTTTTCTTTTCT 4500  
4441 TGAATGACATGAGACCAATGTGTCTCATTTACATTTCTTTTCTTTTCTTTTCT 4500  
4501 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 4560  
4501 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 4560  
4561 TGGCGCAATCTGGGCTCATCTGCAACCTGCTGCTGGGCTCAAGCAATTTCTCCACTC 4620  
4561 TGGCGCAATCTGGGCTCATCTGCAACCTGCTGCTGGGCTCAAGCAATTTCTCCACTC 4620  
4621 AGCTCTCCAAATAGCTGGGATCACTGGCAACCAACATGCCAGCTAATTTTGTATTT 4680  
4621 AGCTCTCCAAATAGCTGGGATCACTGGCAACCAACATGCCAGCTAATTTTGTATTT 4680  
4681 TTTTGTAGACAGAGGTTTACCATTTGTTGCCAGAGCTGTCTCAACTCTGGGCTCAAC 4740  
4681 TTTTGTAGACAGAGGTTTACCATTTGTTGCCAGAGCTGTCTCAACTCTGGGCTCAAC 4740  
4741 AATCTCTGCTGCTGGGCTCCCAAGTGTGCTGGGATTAAGATGTGAGCCAGCATCCAG 4800  
4741 AATCTCTGCTGCTGGGCTCCCAAGTGTGCTGGGATTAAGATGTGAGCCAGCATCCAG 4800  
4801 CCCACACCTCTATTATTAACCAATTAACCTGCTGCTGCTGCTGCTGCTGCTGCTG 4860  
4801 CCCACACCTCTATTATTAACCAATTAACCTGCTGCTGCTGCTGCTGCTGCTGCTG 4860  
4861 CCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4920  
4861 CCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4920  
4921 CAATATTTCTGCGGCAAACTCTCTCAAAATTAATGATTAATGATTAATGATTAATG 4980  
4921 CAATATTTCTGCGGCAAACTCTCTCAAAATTAATGATTAATGATTAATGATTAATG 4980  
4981 ATTCACTTGGAAATGCAACGCTCAACCTGTTCAATGATGATTAATGAAAGAAATTTTA 5040  
4981 ATTCACTTGGAAATGCAACGCTCAACCTGTTCAATGATGATTAATGAAAGAAATTTTA 5040  
5041 TAGTCTCTTAATAGGCTGTAATGCAAGACTCTTGAACACTTTCCAGAGGATAGATAT 5100  
5041 TAGTCTCTTAATAGGCTGTAATGCAAGACTCTTGAACACTTTCCAGAGGATAGATAT 5100  
5101 TTAAGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5160  
5101 TTAAGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5160  
5161 CCAAGTACCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5220  
5161 CCAAGTACCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5220  
5221 CTGCGGCTTCTTCT 5280  
5221 CTGCGGCTTCTTCT 5280  
5281 CCTGCATCTGCAATTTCCCAACCTCCAGGCTTCCAGAGGATAGAGGCTGCTGCTG 5340  
5281 CCTGCATCTGCAATTTCCCAACCTCCAGGCTTCCAGAGGATAGAGGCTGCTGCTG 5340

5341 ATTCACCCAAAGTGGGATTTGGCTCTCTTAGGCTGGCTACTTGTACCATTCACGACAT 5400  
5341 ATTCACCCAAAGTGGGATTTGGCTCTCTTAGGCTGGCTACTTGTACCATTCACGACAT 5400  
5401 CACTGTGCTGCAAGGACACCAAGTGGCAATTTCTTCAACATGAGGCTCAAAATCTC 5460  
5401 CACTGTGCTGCAAGGACACCAAGTGGCAATTTCTTCAACATGAGGCTCAAAATCTC 5460  
5461 TGGACAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5520  
5461 TGGACAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5520  
5521 CCAGCTTGAAGAACCTTGGCTCTTTTCTTTTAAAGCCAGGCTTCAATTAATTAAT 5580  
5521 CCAGCTTGAAGAACCTTGGCTCTTTTCTTTTAAAGCCAGGCTTCAATTAATTAAT 5580  
5581 TGAAGGTACCTGGAACAGTGAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5640  
5581 TGAAGGTACCTGGAACAGTGAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5640  
5641 GCTCATCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5700  
5641 GCTCATCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5700  
5701 AATGCTGTGCTATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5760  
5701 AATGCTGTGCTATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5760  
5761 CAAGCCACCGGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5820  
5761 CAAGCCACCGGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5820  
5821 GCTTCACTGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5880  
5821 GCTTCACTGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5880  
5881 CTGTGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5940  
5881 CTGTGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5940  
5941 CTACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6000  
5941 CTACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6000  
6001 ACATGAGGCTGCAAGAGTGTCTCTTAATTTGTGATGATACAGAAAGCAAGCTTTGGG 6060  
6001 ACATGAGGCTGCAAGAGTGTCTCTTAATTTGTGATGATACAGAAAGCAAGCTTTGGG 6060  
6061 ATACAAATTTCTTCTCTTCAATTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6120  
6061 ATACAAATTTCTTCTCTTCAATTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6120  
6121 GGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6163  
6121 GGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6163

RESULT 2  
US-09-814-353-20489  
Sequence 20489, Application US/09814353  
Publication No. US20030165831A1  
GENERAL INFORMATION:  
APPLICANT: Lee, John  
APPLICANT: Thompson, Pamela  
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
THERAPY OF OVARIAN CANCER  
FILE REFERENCE: MRI-006B  
CURRENT FILING DATE: 2001-03-21  
CURRENT APPLICATION NUMBER: US/09/814,353  
PRIORITY FILING DATE: 2000-03-21  
PRIORITY FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: US 60/207,124  
PRIOR FILING DATE: 2000-05-25  
PRIOR APPLICATION NUMBER: US 60/211,940  
PRIOR FILING DATE: 2000-06-15  
PRIOR APPLICATION NUMBER: US 60/216,820  
PRIOR FILING DATE: 2000-07-07  
PRIOR APPLICATION NUMBER: US 60/220,661  
PRIOR FILING DATE: 2000-07-25  
PRIOR APPLICATION NUMBER: US 60/257,672  
PRIOR FILING DATE: 2000-12-21  
NUMBER OF SEQ ID NOS: 22037  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 20489  
LENGTH: 6286  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 1, 2, 3, 4, 5, 6284, 6285, 6286  
OTHER INFORMATION: n = A,T,C or G  
US-09-814-353-20489

Query Match 87.6%; Score 5400; DB 13; Length 6286;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 5840; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

309 TATCCAGCTGCTAGGGGTTCTGCTGCTGGGCGCGCCCTGCGCGGGGAGAAAC 368  
66 TATCCAGCTGCTAGGGGTTCTGCTGCTGGGCGCGCCCTGCGCGGGGAGAAAC 125  
369 TTTTGAAGTTGCTGCTGCGACGAGAAAGCAATTCAGTTCTCATTAAGCTGGGAGCCC 428  
126 TTTTGAAGTTGCTGCTGCGACGAGAAAGCAATTCAGTTCTCATTAAGCTGGGAGCCC 185  
429 GACTGCTGCTGCGAAACCTGTTTACATGCTCATTTTAAAGACATATACATGTTGTC 488  
186 GACTGCTGCTGCGAAACCTGTTTACATGCTCATTTTAAAGACATATACATGTTGTC 245  
489 CATCAAGCTGAGAAAGAAATAGCTTTTACCTTTAGCTGCGACAGAGCTGAGAACTACCT 548  
246 CATCAAGCTGAGAAAGAAATAGCTTTTACCTTTAGCTGCGACAGAGCTGAGAACTACCT 305  
549 TGTCTAGAGATCCAGAAATAATATGACTGTATGTCAGGCCCATGTCCTTTTGGGAGGT 608  
306 TGTCTAGAGATCCAGAAATAATATGACTGTATGTCAGGCCCATGTCCTTTTGGGAGGT 365  
609 TAGCTTCAAGCTGCACTGCTGTTGCTTACCTTCAACAGAACTTTTATCTGGAGTGT 668  
366 TCAGCTTCAAGCTGCACTGCTGTTGCTTACCTTCAACAGAACTTTTATCTGGAGTGT 425  
669 CAAGCTCATAGAGATGCTTGTAGAGCTGAGTTTTCATCCCTGCGCTGAGGAGAT 728  
426 CAAGCTCATAGAGATGCTTGTAGAGCTGAGTTTTCATCCCTGCGCTGAGGAGAT 484  
729 CGGTCGGGTGAGAGCTGCGACAGAGAGTCACTCATCTCAGCGGCCGAGTGCATGTC 788  
485 CGGTCGGGTGAGAGCTGCGACAGAGAGTCACTCATCTCAGCGGCCGAGTGCATGTC 544  
789 CACCGTGTCAAGATCGAACTCTTCTGACGAAATGAGCACTGTGTCCGGATCAAGATGA 848  
545 CACCGTGTCAAGATCGAACTCTTCTGACGAAATGAGCACTGTGTCCGGATCAAGATGA 604  
849 AGAAGAGTGAAGATGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 908  
605 AGAAGAGTGAAGATGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 664  
909 CAGCATTCGAAACCGCTCATCTATAAAGCTGTGTCATCATGATGCTGTGTTTGAAGG 968  
665 CAGCATTCGAAACCGCTCATCTATAAAGCTGTGTCATCATGATGCTGTGTTTGAAGG 724  
969 TGAAGGCTCAGCAACCTGATGCTGCGCAAGTACCCAGAAAGCTTCCCTGAGAGTACCT 1028  
725 TGAAGGCTCAGCAACCTGATGCTGCGCAAGTACCCAGAAAGCTTCCCTGAGAGTACCT 784

1029 CATGAGTGGCAGTTTGTGTTCTGTCACACCTGGGGCAGAGTCTCTCTCAACTT 1088  
785 CATGAGTGGCAGTTTGTGTTCTGTCACACCTGGGGCAGAGTCTCTCTCAACTT 844  
1089 CAACCTTCTCAACTGTGAGAGAGAGAGAGGAGGAGTGAATCTATCATCTCCGGGCTCCAC 1148  
845 CAACCTTCTCAACTGTGAGAGAGAGAGAGGAGGAGTGAATCTATCATCTCCGGGCTCCAC 904  
1149 CACCAACCCGAGAGTTCAGAGTGGAGAGAGAGAGTGGAGATCTGAGAGAGAGAGTTC 1208  
905 CACCAACCCGAGAGTTCAGAGTGGAGAGAGAGAGTGGAGATCTGAGAGAGAGAGTTC 964  
1209 CAACCTTCTCTGCAAGCTGTCGACCAAGATGCCCCAAGATCTCAGAGATCTCCGCTGCA 1268  
965 CAACCTTCTCTGCAAGCTGTCGACCAAGATGCCCCAAGATCTCAGAGATCTCCGCTGCA 1024  
1269 GTTCCAGATTGTTGTTCCAAATCCAAATGAAAGCAATAAATCTACGTGTTGACTT 1328  
1025 GTTCCAGATTGTTGTTCCAAATCCAAATGAAAGCAATAAATCTACGTGTTGACTT 1084  
1329 GAGTAATGAGAGGAGGAGTCACTCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1388  
1085 GAGTAATGAGAGGAGGAGTCACTCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1144  
1389 GTTGTCCCTGCTGTTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1448  
1145 GTTGTCCCTGCTGTTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1204  
1449 GACATCTGGCTTCCAAACAAATCTCTCTTGTGATGATGATGATGATGATGATGATGAT 1508  
1205 GACATCTGGCTTCCAAACAAATCTCTCTTGTGATGATGATGATGATGATGATGATGAT 1264  
1509 GAATGTGAAAAAACATTAAGCTGACAGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1568  
1265 GAATGTGAAAAAACATTAAGCTGACAGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1324  
1569 ACTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1628  
1325 ACTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1384  
1629 GCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1688  
1385 GCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1444  
1689 GATTAACAACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1748  
1445 GATTAACAACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1504  
1749 CAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1808  
1505 CAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1564  
1809 GAAAGCAGAACTCTGAGTACCTTTCGACCTTTGCGACCTTTGCGACCTTTGCGACCTTT 1868  
1565 GAAAGCAGAACTCTGAGTACCTTTCGACCTTTGCGACCTTTGCGACCTTTGCGACCTTT 1624  
1869 CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1928  
1625 CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1684  
1929 GAGCCCTGACCAAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1988  
1685 GAGCCCTGACCAAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1744  
1989 ATCCCTCACTCTGTGCTCTGAGCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2048  
1745 ATCCCTCACTCTGTGCTCTGAGCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1804  
2049 TTTCTTTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2108  
1805 TTTCTTTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1864

QY 2109 GGAGCAGCGGACCCGGGCTGAGAGATCTTCAAGCTTGAACGAGATGTGCTCCCAAGCC 2168  
DB 1865 GGAGCAGCGGACCCGGGCTGAGAGATCTTCAAGCTTGAACGAGATGTGCTCCCAAGCC 1924  
QY 2169 AAGCTTCCACCATCAAGCTTCTGGGTCAACATCTTAACTGAGGCCACGAGCGGGCAA 2228  
DB 1925 AAGCTTCCACCATCAAGCTTCTGGGTCAACATCTTAACTGAGGCCACGAGCGGGCAA 1984  
QY 2229 GGAGCTAGAAGCTGCTCTTCTGGGTGACCTTACCCCAAGGACTGTGAGCTTGAAGTGCAT 2288  
DB 1985 GGAAGTAAAGCTGCTCTTCTGGGTGACCTTACCCCAAGGACTGTGAGCTTGAAGTGCAT 2044  
QY 2289 CCTCATGAGCGGGTGGAGAGTGAAGTCTTACTGTGCTGCTGCTGGGCTCATCTTTG 2348  
DB 2045 CCTCATGAGCGGGTGGAGAGTGAAGTCTTACTGTGCTGCTGCTGGGCTCATCTTTG 2104  
QY 2349 CTGTGTAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 2408  
DB 2105 CTGTGTAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 2164  
QY 2409 CAACATCAATTAAGATGAGATGCGGAGCGCAAGAAAGTTTCAGAAAGGCGGAAAGACAA 2468  
DB 2165 CAACATCAATTAAGATGAGATGCGGAGCGCAAGAAAGTTTCAGAAAGGCGGAAAGACAA 2224  
QY 2469 TGAATCTCCATGTGTATGAGATCATCGAGGACACCATGTATATGGGCAATCTGTACAGGA 2528  
DB 2225 TGAATCTCCATGTGTATGAGATCATCGAGGACACCATGTATATGGGCAATCTGTACAGGA 2284  
QY 2529 TTCCAGCGGCTCTTCTGAGCGCAGAGGTGACACCTACCGGCGTTCCAGGCGCACCAT 2588  
DB 2285 TTCCAGCGGCTCTTCTGAGCGCAGAGGTGACACCTACCGGCGTTCCAGGCGCACCAT 2344  
QY 2589 GGGGGTCTGTCTCTCCCTCCCAACCAACATATGCTCCAGGCGCCCACTGCAAAAGTTGCG 2648  
DB 2345 GGGGGTCTGTCTCTCCCTCCCAACCAACATATGCTCCAGGCGCCCACTGCAAAAGTTGCG 2404  
QY 2649 CACTGAGAGCCACTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2708  
DB 2405 CACTGAGAGCCACTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2464  
QY 2709 CCATCCCAACATGGGGATGTATGAGCAAGGACACAGACATTCCTTACTGTAACCTCA 2768  
DB 2465 CCATCCCAACATGGGGATGTATGAGCAAGGACACAGACATTCCTTACTGTAACCTCA 2524  
QY 2769 GGAGCCCATGAGCCAGCAAGATTAATTCATTCAGACGCTTGTGCTGATTTTCA 2828  
DB 2525 GGAGCCCATGAGCCAGCAAGATTAATTCATTCAGACGCTTGTGCTGATTTTCA 2584  
QY 2829 AAGCAGGGCACTGAGACACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2888  
DB 2585 AAGCAGGGCACTGAGACACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2644  
QY 2889 ATACAGAAGAACAGAGAGGTTTCTGAGACAGCGCAACTTCACTTGTCACTGGA 2948  
DB 2645 ATACAGAAGAACAGAGAGGTTTCTGAGACAGCGCAACTTCACTTGTCACTGGA 2704  
QY 2949 CTCATTTCTAAGGCGAAGCACTTGAATATGATTAATTCATCTGATTAAGTCACTGAG 3008  
DB 2705 CTCATTTCTAAGGCGAAGCACTTGAATATGATTAATTCATCTGATTAAGTCACTGAG 2764  
QY 3009 CTCATGTGCTCTCAACTTAAGGCTGTGCGGTTAGCCAGCCTGTATATGAGAGAGAGC 3068  
DB 2765 CTCATGTGCTCTCAACTTAAGGCTGTGCGGTTAGCCAGCCTGTATATGAGAGAGAGC 2824  
QY 3069 CTGAGTCACTAGACATAGGTTGACAGAGCCTGTATATGAGAGAGAGAGAGAGAGAG 3128  
DB 2825 CTGAGTCACTAGACATAGGTTGACAGAGCCTGTATATGAGAGAGAGAGAGAGAGAG 2884  
QY 3129 GGCCTTTGAGAGCAACAGTTTCAATTTCAAGAGGCTTACCTGAGTCTCTCACT 3188  
DB 2885 GGCCTTTGAGAGCAACAGTTTCAATTTCAAGAGGCTTACCTGAGTCTCTCACT 2944  
QY 3189 GGGGTCCCAAGATGAAAGCAACATGTGCTTTTATTTATTTATTTGTTGTTGCTG 3248

DB 2945 GGGGTCCCAAGATGAAAGCAACATGTGCTTTTATTTATTTATTTGTTGTTGCTG 3004  
QY 3249 TGTATTTAAGAGATCAATATGATTAACACCTAGCTCTTTTACCTGACTTATTAAC 3308  
DB 3005 TGTATTTAAGAGATCAATATGATTAACACCTAGCTCTTTTACCTGACTTATTAAC 3064  
QY 3309 TCATACATACTGTTTGAATGCTGGTTGACTTCTTACAGCCGTATGATTAACGCTGT 3368  
DB 3065 TCATACATACTGTTTGAATGCTGGTTGACTTCTTACAGCCGTATGATTAACGCTGT 3124  
QY 3365 GGCCTTCCCAAGGAGTGGAGAAATTAATTAACATGTCACCAAGAAAGAGTGTGT 3428  
DB 3125 GGCCTTCCCAAGGAGTGGAGAAATTAATTAACATGTCACCAAGAAAGAGTGTGT 3184  
QY 3429 GTTTAGAGCACTTACACATATCTGCTTTGATTAAGAGCTTCTGATTTCTAGGTCG 3488  
DB 3185 GTTTAGAGCACTTACACATATCTGCTTTGATTAAGAGCTTCTGATTTCTAGGTCG 3244  
QY 3489 TTGTTGTTATCCATTTGTGAAATTCATCTTGAATCCCATTTGCTTATGCTTACAA 3548  
DB 3245 TTGTTGTTATCCATTTGTGAAATTCATCTTGAATCCCATTTGCTTATGCTTACAA 3304  
QY 3549 TTAAGAAATTTCTCAAGTTTCCATGTCGGTTCCTAGCTGAGCAATACCTTTGA 3608  
DB 3305 TTAAGAAATTTCTCAAGTTTCCATGTCGGTTCCTAGCTGAGCAATACCTTTGA 3364  
QY 3609 TTTAAGAGAAATTTAGAGAAATATCTCATCTCTTAAATGTTTAAATATATACCAAC 3668  
DB 3365 TTTAAGAGAAATTTAGAGAAATATCTCATCTCTTAAATGTTTAAATATATACCAAC 3424  
QY 3669 AGTGACCCCTGCAATTAATTTTCTGTTGCACTGCAACCCATTAATCTTGTACTTAA 3728  
DB 3425 AGTGACCCCTGCAATTAATTTTCTGTTGCACTGCAACCCATTAATCTTGTACTTAA 3484  
QY 3729 CAACATTAAGTTTATGCTTGGGGATCAGAAATTCAAATAGATGTCCTGAATGA 3788  
DB 3485 CAACATTAAGTTTATGCTTGGGGATCAGAAATTCAAATAGATGTCCTGAATGA 3544  
QY 3789 ATCAAGTGTGAGCAGAGCTGTGCTCTTCTGAAAGCTTGAAGAGACCGGTTCTT 3848  
DB 3545 ATCAAGTGTGAGCAGAGCTGTGCTCTTCTGAAAGCTTGAAGAGACCGGTTCTT 3604  
QY 3849 CCATTTCAAGCTTCTAAGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3908  
DB 3605 CCATTTCAAGCTTCTAAGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3664  
QY 3909 CTCAATGGCACTAGTGAGACTGGCCCTCCACTTCCCTTGTGACTTACAAGGCCA 3968  
DB 3665 CTCAATGGCACTAGTGAGACTGGCCCTCCACTTCCCTTGTGACTTACAAGGCCA 3724  
QY 3969 CCAGAAAGTCCAGAGTATCTCTCAATCTAAGATCTTCACTATCTCTGAAAGACCTT 4028  
DB 3725 CCAGAAAGTCCAGAGTATCTCTCAATCTAAGATCTTCACTATCTCTGAAAGACCTT 3784  
QY 4029 TTGCATGCAAGCAACATGACAGAGTGGGATTAAGACAGAGACATCTTTGGGCTC 4088  
DB 3785 TTGCATGCAAGCAACATGACAGAGTGGGATTAAGACAGAGACATCTTTGGGCTC 3844  
QY 4089 TGTATTTCTGCTTACCAACTTCTGCACTGATCTCCACAGAGAGGCTCAAAATGA 4148  
DB 3845 TGTATTTCTGCTTACCAACTTCTGCACTGATCTCCACAGAGAGGCTCAAAATGA 3904  
QY 4149 TCTGCGCAGCAGAGATGTTTGTATGCTGCGGACTTAACTT-AAAAAACCAG 4207  
DB 3905 TCTGCGCAGCAGAGATGTTTGTATGCTGCGGACTTAACTT-AAAAAACCAG 3964  
QY 4208 ATCAGAAAGTCTGGCATGCTGGGCTCACTTCACTAGCAACACTGCTGAGCT 4267  
DB 3965 ATCAGAAAGTCTGGCATGCTGGGCTCACTTCACTAGCAACACTGCTGAGCT 4024  
QY 4268 GGGCACAGCTTGCTTTTGAAGGGGTGTCACTTCACTAGCAACACTGCTGAGCT 4327

4025 GGGACCAAGCTTCCCTTTAGAAAGGGGTGTCCATTCAACAGGTCACCAAGCCCAACT 4084  
4328 AGCCCTTATCACTTCCCAATAGAGGCTAAAGTTGTTTCTACTGATCAATGCCCCCTGC 4387  
4085 AGCCCTTATCACTTCCCAATAGAGGCTAAAGTTGTTTCTACTGATCAATGCCCCCTGC 4144  
4388 AGGTTGCAATTTATGTAATGAAAAAGAAAGACTGGAGTTAATCTCTAATCAGGTAGTAG 4447  
4145 AGGTTGCAATTTATGTAATGAAAAAGAAAGACTGGAGTTAATCTCTAATCAGGTAGTAG 4204  
4448 ACCATAGAGCAAAATGAGGCTCAATTAACCTTTTCTTTTCTTTTCTTTTCTTTTCTT 4507  
4205 ACCATAGAGCAATGAGGCTCAATTAACCTTTTCTTTTCTTTTCTTTTCTTTTCTT 4264  
4508 TTTTCTTTTATATGAGACAGAGATCTATTCTGTTGCTAGGCTGAGAGTGAAGGCGCA 4567  
4265 TTTTCTTTTATATGAGACAGAGATCTATTCTGTTGCTAGGCTGAGAGTGAAGGCGCA 4324  
4568 ATCTGGGCTCACTGCAACCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4627  
4325 ATCTGGGCTCACTGCAACCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4384  
4628 CAATAGAGGAGTATCACTGAGCAAAACCAACCTGCTGCTGCTGCTGCTGCTGCTGCTG 4687  
4385 CAATAGAGGAGTATCACTGAGCAAAACCAACCTGCTGCTGCTGCTGCTGCTGCTGCTG 4444  
4688 AGACAGGAGTATCACTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4747  
4445 AGACAGGAGTATCACTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4504  
4748 CTGCTGAGGCTTCCCAAAAGTGTGAGATTAAGATGTAGGCAAGCCGCAATCCAGCCCA 4807  
4505 CTGCTGAGGCTTCCCAAAAGTGTGAGATTAAGATGTAGGCAAGCCGCAATCCAGCCCA 4564  
4808 CCCCATTATTAACCAATTAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4867  
4565 CCCCATTATTAACCAATTAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4624  
4868 CTGATCTGGAAGAGAGGATTAATGTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4927  
4625 CTGATCTGGAAGAGAGGATTAATGTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4684  
4928 TCTGCGGCAAAAACCTTCTTCAAAAATTAATGTAATTAATGTAATTAATGTAATTAAT 4987  
4685 TCTGCGGCAAAAACCTTCTTCAAAAATTAATGTAATTAATGTAATTAATGTAATTAAT 4744  
4988 TTGGAATAGCACCGGCTCACTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5047  
4745 TTGGAATAGCACCGGCTCACTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4804  
5048 CTAAATGAGGAGTATGTAAGAGCTCTTGAACACTTCCAGAGATAGAGATTAATTAAGTC 5107  
4805 CTAAATGAGGAGTATGTAAGAGCTCTTGAACACTTCCAGAGATAGAGATTAATTAAGTC 4864  
5108 ATGCCCTTGGGCTTGGCTTATGAGCACTTTCCTTCTGAAAGCTGCTGCTGCTGCTGCA 5167  
4865 ATGCCCTTGGGCTTGGCTTATGAGCACTTTCCTTCTGAAAGCTGCTGCTGCTGCTGCA 4924  
5168 CCCCCTGCTTGTGAGCGGAGATGCTGCACTGCTGCAATAAGGCGCAAGAGAGGCTGCGGC 5227  
4925 CCCCCTGCTTGTGAGCGGAGATGCTGCACTGCTGCAATAAGGCGCGCAAGAGAGGCTGCGGC 4984  
5228 TTCTCTTCCCTCACTGAGAGGCTTATTTGAATTAAGTGTGAGGCTGCTGCTGCTGCTGCT 5287  
4985 TTCTCTTCCCTCACTGAGAGGCTTATTTGAATTAAGTGTGAGGCTGCTGCTGCTGCTGCT 5044  
5288 TCTGAGACATTTCCCAACCTTCCAGGCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5347  
5045 TCTGAGACATTTCCCAACCTTCCAGGCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5104  
5348 CCAAGGAGGAGATGAGCTTCTTCTTCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5407  
5105 CCAAGGAGGAGATGAGCTTCTTCTTCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5164

5408 GCTTCAAGAGACACCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5467  
5165 GCTTCAAGAGACACCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5224  
5468 GTTGTGCTGCTTCTGAGACAGATTAATTTCTGAGAGCTGCTGCTGCTGCTGCTGCTG 5527  
5225 GTTGTGCTGCTTCTGAGACAGATTAATTTCTGAGAGCTGCTGCTGCTGCTGCTGCTG 5284  
5528 GAGGAGACCTGCTGCTTCTTCTTAAAGCCAGGCCCCCTTAATTAAGATTAAGATTA 5587  
5285 GAGGAGACCTGCTGCTTCTTCTTAAAGCCAGGCCCCCTTAATTAAGATTAAGATTA 5344  
5588 CACTGGAAGACAGTGAAGTGCATTTGTTGAAGCTTAAGCTTAAGCTTAAGCTTAAGCT 5647  
5345 CACTGGAAGACAGTGAAGTGCATTTGTTGAAGCTTAAGCTTAAGCTTAAGCTTAAGCT 5404  
5648 CAGTGTCTGCTGCTGCTTCTGAGAGAGGCTGCTGAGAGAGGCTGCTGCTGCTGCTG 5707  
5405 CAGTGTCTGCTGCTGCTTCTGAGAGAGGCTGCTGAGAGAGGCTGCTGCTGCTGCTG 5464  
5708 TGTGCTATGCAAGAGAGGAGAGAGGCTGCTGAGAGAGAGGCTGCTGAGAGAGAGG 5767  
5465 TGTGCTATGCAAGAGAGGAGAGAGGCTGCTGAGAGAGAGGCTGCTGAGAGAGAGG 5524  
5768 ACCGAGACATGAGCTTGTGTAAGGTTAGCAGACTGCTGCTGCTGCTGCTGCTGCTG 5827  
5525 ACCGAGACATGAGCTTGTGTAAGGTTAGCAGACTGCTGCTGCTGCTGCTGCTGCTG 5584  
5828 TGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5887  
5585 TGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5644  
5888 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5947  
5645 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5704  
5948 TCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6007  
5705 TCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5764  
6008 GCTGAGCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6067  
5765 GCTGAGCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5824  
6068 TTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6127  
5825 TTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5884  
6128 AAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6153  
5885 AAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5910

RESULT 3  
US-09-899-569a-1  
; Sequence 1, Application US/0989569A  
; Patent No. US20020142003A1  
; GENERAL INFORMATION:  
; APPLICANT: No. US20020142003Albert Schweitzer  
; APPLICANT: Marwa Scherl-Mostageer  
; APPLICANT: Wolfgang Sommerhuber  
; APPLICANT: Roger Abscher  
; TITLE OF INVENTION: Tumorasozietres Antigen (B345)  
; FILE REFERENCE: 0652.2280001  
; CURRENT FILING DATE: US/09/899, 569A  
; PRIOR FILING DATE: 2001-07-06  
; PRIOR APPLICATION NUMBER: DE 100 33 080.0  
; PRIOR FILING DATE: 2000-07-07  
; PRIOR APPLICATION NUMBER: DE 101 19 294.0  
; PRIOR FILING DATE: 2001-04-19  
; PRIOR APPLICATION NUMBER: US 60/243,158  
; PRIOR FILING DATE: 2000-10-25



PRIOR APPLICATION NUMBER: US 60/297,747  
 PRIOR FILING DATE: 2001-06-14  
 NUMBER OF SEQ ID NOS: 40  
 SOFTWARE: Patentin Ver. 2.1  
 SEQ ID NO 1  
 LENGTH: 5897  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: misc\_feature  
 LOCATION: (5348)...(5348)  
 OTHER INFORMATION: n is any nucleotide of a, t, g or c  
 NAME/KEY: misc\_feature  
 LOCATION: (5360)...(5360)  
 OTHER INFORMATION: n is any nucleotide of a, t, g or c  
 NAME/KEY: misc\_feature  
 LOCATION: (5425)...(5425)  
 OTHER INFORMATION: n is any nucleotide of a, t, g or c  
 NAME/KEY: misc\_feature  
 LOCATION: (5438)...(5438)  
 OTHER INFORMATION: n is any nucleotide of a, t, g or c  
 NAME/KEY: 5' UTR  
 LOCATION: (1)...(214)  
 NAME/KEY: CDS  
 LOCATION: (215)...(2464)  
 NAME/KEY: 3' UTR  
 LOCATION: (2465)...(5897)  
 US-09-899-569A-1

Query Match 87.5%; Score 5395; DB 10; Length 5897;  
 Best Local Similarity 99.9%; Pred. No. 0;

Matches 5865; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

291 CCTGAATCGGGGCTCTATCGACTGCTAGGGGTTCTGCTGGTGGCGGCGCT 350  
 25 CCGAATCGGGGCTCTATCGACTGCTAGGGGTTCTGCTGGTGGCGGCGCT 84  
 351 GCGGCGGGGCGAAGCTTTGAGATTGCTGCGACAGAAAGCAATTACATTCT 410  
 85 GCGGCGGGGCGAAGCTTTGAGATTGCTGCGACAGAAAGCAATTACATTCT 144  
 411 CATTAAGCTGGGAGACCCGACCTGCTGCGAAACCCTGTACATGTCATTTCTAAAG 470  
 145 CATTAAGCTGGGAGACCCGACCTGCTGCGAAACCCTGTACATGTCATTTCTAAAG 204  
 471 ACATATAACCATGTTCTCATCAAGTCTGGAGAAAGATAGTCTTACCTTTAGCTGCA 530  
 205 ACATATAACCATGTTCTCATCAAGTCTGGAGAAAGATAGTCTTACCTTTAGCTGCA 264  
 531 GAGTCTGGAATCACTTTGTCTATAGATCCAGAAAAATATTGACTGTATGTCAAGGCC 590  
 265 GAGTCTGGAATCACTTTGTCTATAGATCCAGAAAAATATTGACTGTATGTCAAGGCC 324  
 591 ATGTCTTTTGGGAGAGTTCAAGTTCAAGCCCTCGAATCGTTGTCCTTACCTCAAG 650  
 325 ATGTCTTTTGGGAGAGTTCAAGTTCAAGCCCTCGAATCGTTGTCCTTACCTCAAG 384  
 651 AACTTTCATCTGGAGTGTCAAAAGCTCATTAAGAGCATCGTTTGAAGCTCAAGTTTCCAT 710  
 385 AACTTTCATCTGGAGTGTCAAAAGCTCATTAAGAGCATCGTTTGAAGCTCAAGTTTCCAT 444  
 711 CCCTCGCTGAGGCAAGTCGGTCCGGGTGAGAGCTGCCAGAGAGTCACTCATTCAT 770  
 445 CCCTCGCTGAGGCAAGTCGGTCCGGGTGAGAGCTGCCAGAGAGTCACTCATTCAT 504  
 771 CAGCGGCGAATGATGATGCCACCGTGTCAAGATCGGAACCTTTTGCAGCAATGGCACTGT 830  
 505 CAGCGGCGAATGATGATGCCACCGTGTCAAGATCGGAACCTTTTGCAGCAATGGCACTGT 564  
 831 GTCCGGGATCAAGATCAAGAGAGGATGAAGATGGCTTACACTCCCATGTTCCACCC 890  
 565 GTCCGGGATCAAGATCAAGAGAGGATGAAGATGGCTTACACTCCCATGTTCCACCC 624

891 CAGAAATGTCTCCGGCTTACAGATTGCAACCGCTCATCTATAAAGCTGTGTCATCAT 950  
 625 CAGAAATGTCTCCGGCTTACAGATTGCAACCGCTCATCTATAAAGCTGTGTCATCAT 684  
 951 CGAGTCTGTGTTGAGGGTGAAGGCTCAGCAACCTGTATGTCTGCCAATACCCAGAAAG 1010  
 685 CGAGTCTGTGTTGAGGGTGAAGGCTCAGCAACCTGTATGTCTGCCAATACCCAGAAAG 744  
 1011 CTTCCCTAGAGATGAGTCTATGACGTGGCAAGTTTGTCTTCTGCAACCTGCGGCGCAG 1070  
 745 CTTCCCTAGAGATGAGTCTATGACGTGGCAAGTTTGTCTTCTGCAACCTGCGGCGCAG 804  
 1071 CGTCTCTCTTCTCACTTCAACCTCTCAACCTGTGAGAGAAAGAGGCGGGTTGAATA 1130  
 805 CGTCTCTCTTCTCACTTCAACCTCTCAACCTGTGAGAGAAAGAGGCGGGTTGAATA 864  
 1131 CTACATCCCGGGCTCCACCAACCCCGAGGTGTTCAAGCTGAGAGCAAGCAGCTTGG 1190  
 865 CTACATCCCGGGCTCCACCAACCCCGAGGTGTTCAAGCTGAGAGCAAGCAGCTTGG 924  
 1191 GAACATGCGGGGAACTTCAACCTCTCTGCAAGGCTGTGACCAAGATGCCCAAGTCC 1250  
 925 GAACATGCGGGGAACTTCAACCTCTCTGCAAGGCTGTGACCAAGATGCCCAAGTCC 984  
 1251 AGGATCTCTCGGGCTGACGATTCGAAGTTTGGTCCAAATCCAAATGAAGCAATTA 1310  
 985 AGGATCTCTCGGGCTGACGATTCGAAGTTTGGTCCAAATCCAAATGAAGCAATTA 1044  
 1311 AATCTAGCTGTTGACTGAGTAATGAGAGCCATGTCATCACTACGAGCCAGCGCC 1370  
 1045 AATCTAGCTGTTGACTGAGTAATGAGAGCCATGTCATCACTACGAGCCAGCGCC 1104  
 1371 CGTCAACAGAGCGGCAAGTTTGTCTGCGCTGTTTCTGTGTCTAAGATCTCGAAGCTG 1430  
 1105 CGTCAACAGAGCGGCAAGTTTGTCTGCGCTGTTTCTGTGTCTAAGATCTCGAAGCTG 1164  
 1431 CAGTAGCAACCTCAACCTGACATCTGCTCCAAACAAATCTCTCTTGTGATGA 1490  
 1165 CAGTAGCAACCTCAACCTGACATCTGCTCCAAACAAATCTCTCTTGTGATGA 1224  
 1491 TCTGACAGCTGTGTGATGAATGTGGAATAAATCAATTAAGCTGACAGACCAAGGATCTG 1550  
 1225 TCTGACAGCTGTGTGATGAATGTGGAATAAATCAATTAAGCTGACAGACCAAGGATCTG 1284  
 1551 CCAGAGAAATCTTACTCACTTCAAGTCCGAGTACATCTTCACTGCTGTGAGCT 1610  
 1285 CCAGAGAAATCTTACTCACTTCAAGTCCGAGTACATCTTCACTGCTGTGAGCT 1344  
 1611 GCATGACTTCTCTGGAAGCTGTGTGCTCCAGAGACAGGCTCAAGCTGTGTGCTG 1670  
 1345 GCATGACTTCTCTGGAAGCTGTGTGCTCCAGAGACAGGCTCAAGCTGTGTGCTG 1404  
 1671 AGCCAGAGCTGACAGCATACACAGAGAGCCCTGCAACACAGCTTACGACTCT 1730  
 1405 AGCCAGAGCTGACAGCATACACAGAGAGCCCTGCAACACAGCTTACGACTCT 1464  
 1731 CGTGGCCAGTGCATACCCAGCAGAGACTGTATCTTGTGCTCTTCTGCGGAGGCTC 1790  
 1465 CGTGGCCAGTGCATACCCAGCAGAGACTGTATCTTGTGCTCTTCTGCGGAGGCTC 1524  
 1791 TATCAAGAGATCCAGGTGAGAGCAAAATCTCGGTGACCTTGGCACTTTGGCCCCAG 1850  
 1525 TATCAAGAGATCCAGGTGAGAGCAAAATCTCGGTGACCTTGGCACTTTGGCCCCAG 1584  
 1851 CTTCAACAAGAGGCTCCAGGCAAGGCTGTGCGGTCTTATATCTTATTTCAAGA 1910  
 1585 CTTCAACAAGAGGCTCCAGGCAAGGCTGTGCGGTCTTATATCTTATTTCAAGA 1644  
 1911 GGAAGGCGTTTTCAGGATGACCTTGAACAACAAAGGCTTACTGTGAGACCCCA 1970  
 1645 GGAAGGCGTTTTCAGGATGACCTTGAACAACAAAGGCTTACTGTGAGACCCCA 1704  
 1971 CTGGAGCGGGGCTGCTCATCTCTGTGTCTGGAACATCAGCTGCCAGAGA 2030

Db 1705 CTGGGACCGGGGCTGCCATCCCTCTGTGCTTGAACATCAGGTCGCCAGAA 1764  
Qy 2031 CCAGGTGGCTGCTGACTTTCTTTAAGAGCGGAGCGGTGCTGCGACAGGGCG 2090  
Db 1765 CCAAGTGGCTGCTGACTTTCTTTAAGAGCGGAGCGGTGCTGCGACAGGGCG 1824  
Qy 2091 CGCATTGATGATCATCAGAGACAGCGGACCGGGCTGAGAGATTTCCAGCTGAGCA 2150  
Db 1825 CGCATTGATGATCATCAGAGACAGCGGACCGGGCTGAGAGATTTCCAGCTGAGCA 1884  
Qy 2151 GGATGTGCTCCCAAGCCAGCTTCCACATCAGAGCTTGTGGGTCAATCTTACTG 2210  
Db 1885 GGATGTGCTCCCAAGCCAGCTTCCACATCAGAGCTTGTGGGTCAATCTTACTG 1944  
Qy 2211 CAGCCCAAGAGCGGCAAGCTAGACCTGCTCTTGTGGTGAACAATTACCCAAAGAC 2270  
Db 1995 CAGCCCAAGAGCGGCAAGCTAGACCTGCTCTTGTGGTGAACAATTACCCAAAGAC 2004  
Qy 2271 TGTGACTTGTATCTCTCATCGACGGTGGAGGTGAGTTTACTGCTGTCTG 2330  
Db 2005 TGTGACTTGTATCTCTCATCGACGGTGGAGGTGAGTTTACTGCTGTCTG 2064  
Qy 2331 CCTGGGCTCATCATTTGCTGTGTGAAAAAGAAAAAGAAACAAGGGCCCGC 2390  
Db 2065 CCTGGGCTCATCATTTGCTGTGTGAAAAAGAAAAAGAAACAAGGGCCCGC 2124  
Qy 2391 TGTGGGTATCTCAATGGCAATCAATCTAGATGGCGAGCGCAAAAAAGTTCA 2450  
Db 2125 TGTGGGTATCTCAATGGCAATCAATCTAGATGGCGAGCAAAAAAGTTTCA 2183  
Qy 2451 GAAAGGCGCAAGAGCAATGATCTCCATGTATGAGTCAATGAGACACCATGATATA 2510  
Db 2184 GAAAGGCGCAAGAGCAATGATCTCCATGTATGAGTCAATGAGACACCATGATATA 2243  
Qy 2511 TGGGCACTGTCTACAGAGATTCAGCGGGCTCTTCTGCAAGCAAGTGGACACTTAC 2570  
Db 2244 TGGGCACTGTCTACAGAGATTCAGCGGGCTCTTCTGCAAGCAAGTGGACACTTAC 2303  
Qy 2571 GCGGTTCAGGGGACCATGGGGGTGTGCTCTCCGCCCAACCATATGCTCCAGGGC 2630  
Db 2304 GCGGTTCAGGGGACCATGGGGGTGTGCTCTCCGCCCAACCATATGCTCCAGGGC 2363  
Qy 2631 CCCAATGCAAGATTGGCACTGAGGAGCACCTCTGCTCCCTCTGAGTGTAGAG 2690  
Db 2364 CCCAATGCAAGATTGGCACTGAGGAGCACCTCTGCTCCCTCTGAGTGTAGAG 2423  
Qy 2691 TGAACCGTACACCTTCTCCCATCCCAACAATGGGGATGTAGCAGCAAGACACAT 2750  
Db 2424 TGAACCGTACACCTTCTCCCATCCCAACAATGGGGATGTAGCAGCAAGACACAT 2483  
Qy 2751 TCCCTTACTGAACACTCAGAGAGCCCATGAGCCAGAGAAATTAATTGATTCAGAC 2810  
Db 2484 TCCCTTACTGAACACTCAGAGAGCCCATGAGCCAGAGAAATTAATTGATTCAGAC 2543  
Qy 2811 GCTTGTCTGAGTTTCTTAAGAGAGGGCACTGAGACACCCGTCCGTGTTCTTAACAGAA 2870  
Db 2544 GCTTGTCTGAGTTTCTTAAGAGAGGGCACTGAGACACCCGTCCGTGTTCTTAACAGAA 2603  
Qy 2871 TCTTAAAGAGAGAAATTAACAGAGAAACAAGAGAGTTTCTGTGACACCGCAAC 2930  
Db 2604 TCTTAAAGAGAGAAATTAACAGAGAAACAAGAGAGTTTCTGTGACACCGCAAC 2663  
Qy 2931 TTCAATTTGCTAGTGAATCTATTTCTAAGGCAAGACATTTGAATGATGATTCATC 2990  
Db 2664 TTCAATTTGCTAGTGAATCTATTTCTAAGGCAAGACATTTGAATGATGATTCATC 2723  
Qy 2991 TGGATACAGTCAATGACAGTCAATGTGCTCTCAACTTAGAGTGTGGGTTAGCAGCCG 3050  
Db 2724 TGGATACAGTCAATGACAGTCAATGTGCTCTCAACTTAGAGTGTGGGTTAGCAGCCG 2783  
Qy 3051 TAATGAGAGAGAGGCTGAGTCACTTAGATAGGTTTGGAGCAAGCCCTGGAATTCA 3110

Db 2784 TAATGAGAGAGAGGCTGAGTCACTTAGATAGGTTTGGAGCAAGCCCTGATTCAG 2843  
Qy 3111 AGTGTAAACAGAGGCTTGCCCTCTTCAGAGCAACAGTTTCAATTCCAGAGGCTTACT 3170  
Db 2844 AGTGTAAACAGAGGCTTGCCCTCTTCAGAGCAACAGTTTCAATTCCAGAGGCTTACT 2903  
Qy 3171 GAGGTCCCTACTCTGACTGGGGTCCCAAGATGAAAAAGACATGTGCTTTTATTAAT 3230  
Db 2904 GAGGTCCCTACTCTGACTGGGGTCCCAAGATGAAAAAGACATGTGCTTTTATTAAT 2963  
Qy 3231 ATTTATTTGGTGTGCTGTGTTATTTAAGATCAATGATTAACCACTTACTTTTC 3290  
Db 2964 ATTTATTTGGTGTGCTGTGTTATTTAAGATCAATGATTAACCACTTACTTTTC 3023  
Qy 3291 ACCGACTTAGATTAATGACTTACTTACCTGTTGGATGGCCGGGTGTTGACTTACTG 3350  
Db 3024 ACCGACTTAGATTAATGACTTACTTACCTGTTGGATGGCCGGGTGTTGACTTACTG 3083  
Qy 3351 ACCGACTTAGATTAATGACTTACTTACCTGTTGGATGGCCGGGTGTTGACTTACTG 3410  
Db 3084 ACCGACTTAGATTAATGACTTACTTACCTGTTGGATGGCCGGGTGTTGACTTACTG 3143  
Qy 3411 CCAGAAAAAGATGTGTGTTTGAAGACATTTGACACATATCTGCTTGAATGAGACTT 3470  
Db 3144 CCAGAAAAAGATGTGTGTTTGAAGACATTTGACACATATCTGCTTGAATGAGACTT 3203  
Qy 3471 CCTGATTTCTTAGGTCGGTGTGGTTTATCCATTGTGGAATTCATCTTGAATCCCAT 3530  
Db 3204 CCTGATTTCTTAGGTCGGTGTGGTTTATCCATTGTGGAATTCATCTTGAATCCCAT 3263  
Qy 3531 GTCTATAGTCTCTAGCAATTAAGAAATTTCTCAAGTTTCCATGTGGGTTCTCTAGC 3590  
Db 3264 GTCTATAGTCTCTAGCAATTAAGAAATTTCTCAAGTTTCCATGTGGGTTCTCTAGC 3323  
Qy 3591 TGCAGCAATCTTGAACATTTAAGAGAAATTTAAGAAATTTCAATCTCTTAAATG 3650  
Db 3324 TGCAGCAATCTTGAACATTTAAGAGAAATTTAAGAAATTTCAATCTCTTAAATG 3383  
Qy 3651 TTTAATATATACAAACAGTGGCCCGCTGACTTGTGTTGCTGCACTGCAACCAT 3710  
Db 3384 TTTAATATATACAAACAGTGGCCCGCTGACTTGTGTTGCTGCACTGCAACCAT 3443  
Qy 3711 TACTGTAGTCTTAAACCAACACATTAAGTATAGTCTGAGGATTCAGAAATTCAGAAAT 3770  
Db 3444 TACTGTAGTCTTAAACCAACACATTAAGTATAGTCTGAGGATTCAGAAATTCAGAAAT 3503  
Qy 3771 GGATGTCCCTGATGATAAATCAAGGTGTGAGCAGAGCTGTGCTCTGAAAGGCTTAG 3830  
Db 3504 GGATGTCCCTGATGATAAATCAAGGTGTGAGCAGAGCTGTGCTCTGAAAGGCTTAG 3563  
Qy 3831 GGAGAGCGCGTTCCTTGCATTTTCAAGCTTCTAAGGCTGCTGATTTCCAGGCTCCA 3890  
Db 3564 GGAGAGCGCGTTCCTTGCATTTTCAAGCTTCTAAGGCTGCTGATTTCCAGGCTCCA 3623  
Qy 3891 GTGGCTGCTCAAGCTTTTCTCATGATGACACTGTGACACTGGGCTCCCACTTCCCTC 3950  
Db 3624 GTGGCTGCTCAAGCTTTTCTCATGATGACACTGTGACACTGGGCTCCCACTTCCCTC 3683  
Qy 3951 TTTGACTTAACAAAGCCCAACAGAAAGATTCAGATTAATCTCTCATTAAGATCTTCA 4010  
Db 3684 TTTGACTTAACAAAGCCCAACAGAAAGATTCAGATTAATCTCTCATTAAGATCTTCA 3743  
Qy 4011 TCAATCTGGAAGAGCTTTTGCATGCAAGACAAATAGCCACAGTGGGATTTAGAAC 4070  
Db 3744 TCAATCTGGAAGAGCTTTTGCATGCAAGACAAATAGCCACAGTGGGATTTAGAAC 3803  
Qy 4071 AGGACATCTTTGGGGTGTGTTATTTCTGCTTACCAACCTTCTGCACTGACTCCACA 4130  
Db 3804 AGGACATCTTTGGGGTGTGTTATTTCTGCTTACCAACCTTCTGCACTGACTCCACA 3863  
Qy 4131 GGAGAGGCTAACAAATGATCTGGCGCACAGGATGTTTGTGTTAGCTTGGGACTTAAAC 4190  
Db 3864 GGAGAGGCTAACAAATGATCTGGCGCACAGGATGTTTGTGTTAGCTTGGGACTTAAAC 3923

QY 4191 ACTTAAAAAACCAGATCAGAAATCTGGCCATCTGGGGCTCACTTTCACCTAGC 4250  
DB 3924 ACTTAAAAAACCAGATCAGAAATCTGGCCATCTGGGGCTCACTTTCACCTAGC 3983  
QY 4251 AACAACTGGGTGGAGCTGGGCAACAGCTGCTGCTTTAGAAAGGGGTGTCCACTTCACAGG 4310  
DB 3984 AACAACTGGGTGGAGCTGGGCAACAGCTGCTGCTTTAGAAAGGGGTGTCCACTTCACAGG 4043  
QY 4311 TCACCAAGCCCACTAGAGCCCTTATCACTTCCCAATAGAGGCTAAGTGTGTTCTA 4370  
DB 4044 TCACCAAGCCCACTAGAGCCCTTATCACTTCCCAATAGAGGCTAAGTGTGTTCTA 4103  
QY 4371 CTGATCAATGCCCTGAGGTGCTGATTTATTTAATGAAAAAGAAAGACTGGGATTAATC 4430  
DB 4104 CTGATCAATGCCCTGAGGTGCTGATTTATTTAATGAAAAAGAAAGACTGGGATTAATC 4163  
QY 4431 TCTAATCAGGTGATGACATAGACCAATGTGTCTCACTTACCCTTTTCTTTT 4490  
DB 4164 TCTAATCAGGTGATGACATAGACCAATGTGTCTCACTTACCCTTTTCTTTT 4223  
QY 4491 TTCTTTTCTTTTCTTTTCTTTTATGTAGAGAGATCTCATTTCTGCTAGGC 4550  
DB 4224 TTCTTTTCTTTTCTTTTCTTTTATGTAGAGAGATCTCATTTCTGCTAGGC 4283  
QY 4551 TGGAGTGCAGTGGCGCAATCTGGCTCACTGCAACTCTGCTCTGGGCTCAAGCAAT 4610  
DB 4284 TGGAGTGCAGTGGCGCAATCTGGCTCACTGCAACTCTGCTCTGGGCTCAAGCAAT 4343  
QY 4611 CTCCCACTCAGGCTCCCAATATGCTGGGATCACTGGCAAAACCAACATGCCACTAA 4670  
DB 4344 CTCCCACTCAGGCTCCCAATATGCTGGGATCACTGGCAAAACCAACATGCCACTAA 4403  
QY 4671 TTTTGTATTTTGTAGAGAGAGGTTTCACTGTTGCCAGGCTGTTCAACTCTCT 4730  
DB 4404 TTTTGTATTTTGTAGAGAGAGGTTTCACTGTTGCCAGGCTGTTCAACTCTCT 4463  
QY 4731 GGGCTCAAGCAATCTCTGCTGCTGCTGCCAAAGTCTGGGATTAAGAGTGAACA 4790  
DB 4464 GGGCTCAAGCAATCTCTGCTGCTGCTGCCAAAGTCTGGGATTAAGAGTGAACA 4523  
QY 4791 CCGCATCCAGCCCAACACCTCATTTATACATTAACCTGCCCAAGTAACGTGGAATTT 4850  
DB 4524 CCGCATCCAGCCCAACACCTCATTTATACATTAACCTGCCCAAGTAACGTGGAATTT 4583  
QY 4851 GCTTCTCAACCTGCTGCTGATCTGGAAGAGAGGATTAATGTATAGCTTGCACACA 4910  
DB 4584 GCTTCTCAACCTGCTGCTGATCTGGAAGAGAGGATTAATGTATAGCTTGCACACA 4643  
QY 4911 GTCCCAAGTTCAATATTTCTGGGCAAAAATTCTTTCAAAAAATATATGATCTTATG 4970  
DB 4644 GTCCCAAGTTCAATATTTCTGGGCAAAAATTCTTTCAAAAAATATATGATCTTATG 4703  
QY 4971 TATTCATGAATTCACCTTGGAAATGACACCGCTCAACTGTTCAATAGGCAATATGA 5030  
DB 4704 TATTCATGAATTCACCTTGGAAATGACACCGCTCAACTGTTCAATAGGCAATATGA 4763  
QY 5031 AGGAATTTTATAGTCTCTAATAGGCTGATCTGAAGACTCTTGAACACTTCCAGAG 5090  
DB 4764 AGGAATTTTATAGTCTCTAATAGGCTGATCTGAAGACTCTTGAACACTTCCAGAG 4823  
QY 5091 GATAGATATTTAATGATGATGCTTGGGCTTGGCTTATGGAACCTTTCTCTGAAAGTC 5150  
DB 4824 GATAGATATTTAATGATGATGCTTGGGCTTGGCTTATGGAACCTTTCTCTGAAAGTC 4883  
QY 5151 TGGTTCTGCGCCAGTACACCTTGGCTTGGAGCCAGATGCTGACACCTTGCATTAAGGGC 5210  
DB 4884 TGGTTCTGCGCCAGTACACCTTGGCTTGGAGCCAGATGCTGACACCTTGCATTAAGGGC 4943  
QY 5211 CAAAGAGAGGCTGCTGCTTCTCTCACTGAAGAGCCCTTATTTGAATTCAGTGTG 5270  
DB 4944 CAAAGAGAGGCTGCTGCTTCTCTCACTGAAGAGCCCTTATTTGAATTCAGTGTG 5003

QY 5271 GAGCCCTAGCCCTCCATTTCTGACATTTGCCCAACTCCAGCCCTTCCAAAGAGACTA 5330  
DB 5004 GAGCCCTAGCCCTCCATTTCTGACATTTGCCCAACTCCAGCCCTTCCAAAGAGACTA 5063  
QY 5331 GGTGCTCTGATTTCCACCAGAGTGGAGTTGGCTTCTCTAGAGCTGGCTACTTGCACA 5390  
DB 5064 GGTGCTCTGATTTCCACCAGAGTGGAGTTGGCTTCTCTAGAGCTGGCTACTTGCACA 5123  
QY 5391 TCACCAATCTCTGAGCAAGTGTGCTGCTGAGACCAATTTCTGAGCTGTGCTCA 5450  
DB 5124 TCACCAATCTCTGAGCAAGTGTGCTGAGACCAATTTCTGAGCTGTGCTCA 5183  
QY 5451 TCAAAATCTCTGAGCAAGTGTGCTGCTGAGACCAATTTCTGAGCTGTGCTCA 5510  
DB 5184 TCAAAATCTCTGAGCAAGTGTGCTGCTGAGACCAATTTCTGAGCTGTGCTCA 5243  
QY 5511 GTAAAGGGGCGAGCTGAGAACCTGGCTTTTCTTTAAAGCCAGGCCCACTTAC 5570  
DB 5244 GTAAAGGGGCGAGCTGAGAACCTGGCTTTTCTTTAAAGCCAGGCCCACTTAC 5303  
QY 5571 ATAAAACATTTCAAGGCTCACTGGAACAGTGAAGTCCATTTGTTGAAGCTTACATG 5630  
DB 5304 ATAAAACATTTCAAGGCTCACTGGAACAGTGAAGTCCATTTGTTGAAGCTTACATG 5363  
QY 5631 CAGCCCACTGCTCATTTCCAGTGTGCTGCTGATGCTTACAGAGAGGCGCAGCTAGCAG 5690  
DB 5364 CAGCCCACTGCTCATTTCCAGTGTGCTGCTGATGCTTACAGAGAGGCGCAGCTAGCAG 5423  
QY 5691 ACTGGCTCTAATGCTGAGTGTGCTTATGCAAGAAAGGAAAGCTGCAAGAGAGTCACT 5750  
DB 5424 ANTGGCTCTAATGCTGAGTGTGCTTATGCAAGAAAGGAAAGCTGCAAGAGAGTCACT 5483  
QY 5751 GGAACAAGCAAGGCCCAACCGGACATGGCTTGTGTAAGGTTAGCAGACTGGTGTGTG 5810  
DB 5484 GGAACAAGCAAGGCCCAACCGGACATGGCTTGTGTAAGGTTAGCAGACTGGTGTGTG 5543  
QY 5811 GATTCGAGTGTCTTCACTGGAATAATTTATTCATTTGCAAGTACTTTTAAAGTGCATTT 5870  
DB 5544 GATTCGAGTGTCTTCACTGGAATAATTTATTCATTTGCAAGTACTTTTAAAGTGCATTT 5603  
QY 5871 TATTCATTTCTGCTGCTTTAAATAAACAATGATACCAAAAACAAGTATCAAGCTGTTA 5930  
DB 5604 TATTCATTTCTGCTGCTTTAAATAAACAATGATACCAAAAACAAGTATCAAGCTGTTA 5663  
QY 5931 AGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5990  
DB 5664 AGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5723  
QY 5991 TGAAGGCTCAGCATGGGCTCAGCAGATGCTGCTTATTTGTGATGATACAGAAAGCC 6050  
DB 5724 TGAAGGCTCAGCATGGGCTCAGCAGATGCTGCTTATTTGTGATGATACAGAAAGCC 5783  
QY 6051 AGGCTTTGGGATACAAAGTTCTTCTCTTCACTTGAATGCGGCACTGTGTAGAGCAT 6110  
DB 5784 AGGCTTTGGGATACAAAGTTCTTCTCTTCACTTGAATGCGGCACTGTGTAGAGCAT 5843  
QY 6111 GTTTTGTCCGGAATAAATAATATAGTCTTGGAGTCTGCCAAAAAATTTT 6163  
DB 5844 GTTTTGTCCGGAATAAATAATATAGTCTTGGAGTCTGCCAAAAAATTTT 5896

RESULT 4  
US-10-128-692A-161  
; Sequence 161, Application US/10128692A  
; Publication No. US20040009547A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Demoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerltzen, Mary E.

APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECURED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P330R1C124  
CURRENT FILING DATE: 2002-10-15  
PRIOR APPLICATION NUMBER: 60/049911  
PRIOR FILING DATE: 1997-06-18  
PRIOR APPLICATION NUMBER: 60/056974  
PRIOR FILING DATE: 1997-08-26  
PRIOR APPLICATION NUMBER: 60/059113  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059115  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059117  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059122  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059184  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059263  
PRIOR FILING DATE: 1997-09-18  
PRIOR APPLICATION NUMBER: 60/059352  
PRIOR FILING DATE: 1997-09-19  
PRIOR APPLICATION NUMBER: 60/059588  
PRIOR FILING DATE: 1997-09-19  
Remaining Prior Application data removed - See File Wrapper or PALM.  
SEQ ID NO 161  
LENGTH: 1376  
TYPE: DNA  
ORGANISM: Homo Sapien  
US-10-128-692A-161

Query Match 18.1%; Score 1117; DB 12; Length 1376;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

190 CCAGGCGCGAGCGCAGCTGCGCGCGGCTTGGGCGCGCTGGGCGCGCTCCCGACCGTGC 249  
1 CCAGGCGCGAGCGCAGCTGCGCGCGGCTTGGGCGCGCTGGGCGCGCTCCCGACCGTGC 60  
250 TTTTCCCAACCGAGCGCGGCTCCCGAGTCATGCGCGGCTGAACTGCGGGCTCTT 309  
61 TTTTCCCAACCGAGCGCGGCTCCCGAGTCATGCGCGGCTGAACTGCGGGCTCTT 120  
310 ATGCACTGCTAGGGGTTTGTCTGCTGAGGCGCGCGCTCCCGCGCGGCGAGAGCT 369  
121 ATGCACTGCTAGGGGTTTGTCTGCTGAGGCGCGCGCTCCCGCGCGGCGAGAGCT 180  
370 TTTGAGATTGCTCTGCGCAGAGAAAGCAATTACAGTTCTCATTAAGCTGGGGAACCCG 429  
181 TTTGAGATTGCTCTGCGCAGAGAAAGCAATTACAGTTCTCATTAAGCTGGGGAACCCG 240  
430 ACTGCGCTGGCAAAACCCCTGTTACATCGCATTTCTAAAGACATATAACCATGTTGTC 489  
241 ACTGCGCTGGCAAAACCCCTGTTACATCGCATTTCTAAAGACATATAACCATGTTGTC 300  
490 ATCAAGTCTGAGAGAAATAGTCTTTACCTTAGCTGCGAGAGTCTGAGAGTCACTTT 549  
301 ATCAAGTCTGAGAGAAATAGTCTTTACCTTAGCTGCGAGAGTCTGAGAGTCACTTT 360  
550 GTCATAGAGATCCAGAAAAATATTAAGCTGTATGTCAGGCCCATGTCCTTTGGGAGGTT 609

361 GTCATAGAGATCCAGAAAAATATTAAGCTGTATGTCAGGCCCATGTCCTTTGGGAGGTT 420  
610 CAGCTTCAAGCCCTGACATCGTTGTTGCTACCCCTCAACAGAACTTTCATCTGGAGTGC 669  
421 CAGCTTCAAGCCCTGACATCGTTGTTGCTACCCCTCAACAGAACTTTCATCTGGAGTGC 480  
670 AAAGCTCATTAAGAGCATCGGTTTAAAGCTGACAGTTTTCATCCCTGCGCTGAGGAGATC 729  
481 AAAGCTCATTAAGAGCATCGGTTTAAAGCTGACAGTTTTCATCCCTGCGCTGAGGAGATC 540  
730 GGTCCGGGTGAGAGTGGCCAGACGAGATGCTACTCTCATATAGAGGCGGCAATGCATGC 789  
541 GGTCCGGGTGAGAGTGGCCAGACGAGATGCTACTCTCATATAGAGGCGGCAATGCATGC 600  
790 ACCGTGTCAGAGTGAACCTTCTGACAGCAATGCACTGTGTCCCGGATCAAGATGCA 849  
601 ACCGTGTCAGAGTGAACCTTCTGACAGCAATGCACTGTGTCCCGGATCAAGATGCA 660  
850 GAAGGAGTGAATAATGCGCTTACACCTCCCATGCTTCAACCCGAAATGTCTCCGCTTC 909  
661 GAAGGAGTGAATAATGCGCTTACACCTCCCATGCTTCAACCCGAAATGTCTCCGCTTC 720  
910 AGCATTTGCAAAACCGCTCATCTATAAAGCTCTGTGATCATGATGATCTGTGAGAGGT 969  
721 AGCATTTGCAAAACCGCTCATCTATAAAGCTCTGTGATCATGATGATCTGTGAGAGGT 780  
970 GAAGGCTCAGCAACCTGATGTCGCAACTACCAAGAGGCTTCCCTGAGAGTGAAGCTC 1029  
781 GAAGGCTCAGCAACCTGATGTCGCAACTACCAAGAGGCTTCCCTGAGAGTGAAGCTC 840  
1030 ATGACGTGGAGTTTGTGCTTCTGACACCTGCGGCGAGCGTCTCTCTCAACTTC 1089  
841 ATGACGTGGAGTTTGTGCTTCTGACACCTGCGGCGAGCGTCTCTCTCTCAACTTC 900  
1090 AACCTCTCCAACTGTGAGAGAGAGAGAGCGGGTTGAATACTACATCCCGGCTCCAC 1149  
901 AACCTCTCCAACTGTGAGAGAGAGAGAGCGGGTTGAATACTACATCCCGGCTCCAC 960  
1150 ACCAACCCCGAGGTTTGAAGCTGAGAGCAAGCAGCTCGGGAACATGCGCGGGAACCTTC 1209  
961 ACCAACCCCGAGGTTTGAAGCTGAGAGCAAGCAGCTCGGGAACATGCGCGGGAACCTTC 1020  
1210 AACCTCTCTGCAAGGCTGTGACCAAGATGCCCAAGTCCAGGATCTCTCGGCTGAG 1269  
1021 AACCTCTCTGCAAGGCTGTGACCAAGATGCCCAAGTCCAGGATCTCTCGGCTGAG 1080  
1270 TTCCAAGTTTGTGTCACATCCACAAATGAAAGCA 1306  
1081 TTCCAAGTTTGTGTCACATCCACAAATGAAAGCA 1117

RESULT 5  
US-10-140-927-161  
Sequence 161, Application US/10140927  
Publication No. US20040009548A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: Deforge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin



```
Db 181 TTGAGATTGCTGTGCACGAGAAAGCAATTACAGTTCTCATAAAGTGGGACCCCG 240
Qy 430 ACTCTGCTGGCAAAACCCGTGTACATCGTCATTTCTAAAGAATATAACATGTTGTCC 489
Db 241 ACTCTGCTGGCAAAACCCGTGTACATCGTCATTTCTAAAGAATATAACATGTTGTCC 300
Qy 490 ATCAAGTCTGGAGAAAGAAATAGTCTTTACCTTTAGCTGCGAGAGTCTTGAAATCACTTT 549
Db 301 ATCAAGTCTGGAGAAAGAAATAGTCTTTACCTTTAGCTGCGAGAGTCTTGAAATCACTTT 360
Qy 550 GTCATATAGATCCAGAAAAATATATGACTGTATGTCAAGGCCCATGTCTTTGGGAGGTT 609
Db 361 GTCATATAGATCCAGAAAAATATATGACTGTATGTCAAGGCCCATGTCTTTGGGAGGTT 420
Qy 610 CAGCTTCAGCCCTCGACATCGTTGTTGGCTTACCCCTCAACAGAACTTTCACTGGAGATGC 669
Db 421 CAGCTTCAGCCCTCGACATCGTTGTTGGCTTACCCCTCAACAGAACTTTCACTGGAGATGC 480
Qy 670 AAAGCTATAGAGCATCGGTTTAAAGCTGAGCTGAGTTTCCATCCCTGCGCTGAGGAGATC 729
Db 481 AAAGCTATAGAGCATCGGTTTAAAGCTGAGCTGAGTTTCCATCCCTGCGCTGAGGAGATC 540
Qy 730 GGTCCGGGTGAGAGCTGCCCAAGCGAGTCACTCACTCATCAAGCGCGCAATCGATGCC 789
Db 541 GGTCCGGGTGAGAGCTGCCCAAGCGAGTCACTCACTCATCAAGCGCGCAATCGATGCC 600
Qy 790 ACCGTGTGAGAGTGGGAACTTTCTGACAGCAATGGACATGTGTCCCGGATCAAGATGCA 849
Db 601 ACCGTGTGAGAGTGGGAACTTTCTGACAGCAATGGACATGTGTCCCGGATCAAGATGCA 660
Qy 850 GAAGAGTGAATAATGAGCTTTACACTCCGATGTTCCACCCAGAAATGTTCTCCGCTTC 909
Db 661 GAAGAGTGAATAATGAGCTTTACACTCCGATGTTCCACCCAGAAATGTTCTCCGCTTC 720
Qy 910 AGCATTTGCAAAACCGCTCATCTATATAAAGCTGTGTGATCATGAGTGTGTGAGGAT 969
Db 721 AGCATTTGCAAAACCGCTCATCTATATAAAGCTGTGTGATCATGAGTGTGTGAGGAT 780
Qy 970 GAAGGCTGAGCAACCTGATGTCTGCAACTACCCAGAAAGCTTCCCTAGAGATGAGCTC 1029
Db 781 GAAGGCTGAGCAACCTGATGTCTGCAACTACCCAGAAAGCTTCCCTAGAGATGAGCTC 840
Qy 1030 ATGACGTGAGCAATTTGCTCTCTGCAACCTGCGGGGAGGCTCTCTTCTCAACTTC 1089
Db 841 ATGACGTGAGCAATTTGCTCTCTGCAACCTGCGGGGAGGCTCTCTTCTCAACTTC 900
Qy 1090 AACCTCTCCAACTGTGAGAGAGAGAGAGCGGGTTGAATACTACAATCCCGGAGCTCCACC 1149
Db 901 AACCTCTCCAACTGTGAGAGAGAGAGAGCGGGTTGAATACTACAATCCCGGAGCTCCACC 960
Qy 1150 ACCAACCCCGAGAGTGTTCAGAGTGAAGAGCAAGACAGCTGGGAAATGAGCGGGAACCTTC 1209
Db 961 ACCAACCCCGAGAGTGTTCAGAGTGAAGAGCAAGACAGCTGGGAAATGAGCGGGAACCTTC 1020
Qy 1210 AACCTCTCTCTGAGAGAGTGTGACCAAGATGCCCAAGTCCAGGAGTCTCTCCGGCTGAG 1269
Db 1021 AACCTCTCTCTGAGAGAGTGTGACCAAGATGCCCAAGTCCAGGAGTCTCTCCGGCTGAG 1080
Qy 1270 TTCCAAGTTTGGTCCAACTCCAAATGAAGA 1306
Db 1081 TTCCAAGTTTGGTCCAACTCCAAATGAAGA 1117
```

RESULT 7  
US-10-140-018-161

; Sequence 161, Application US/10140018  
; Publication No. US20030138885A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc

```
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerilsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P330R1C158
CURRENT APPLICATION NUMBER: US/10/140,018
CURRENT FILING DATE: 2002-05-06
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 161
LENGTH: 1376
TYPE: DNA
ORGANISM: Homo Sapien
US-10-140-018-161

Query Match      18.1%; Score 1117; DB 13; Length 1376;
Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;
Matches 1117; Conservative 0;

Qy 190 CCAGGGCGGAGGCGAGCTGCGCGCGGCTTGGGCGCTTGGGCGCGGCTTCCCAACCGTCG 249
Db 1 CCAGGGCGGAGGCGAGCTGCGCGCGGCTTGGGCGCTTGGGCGCGGCTTCCCAACCGTCG 60
Qy 250 TTTTCCCAACCGAGGCGAGGCGCTTCCGAGTCAATGCGCGGCTTCAATCTGCGGGTCTCT 309
Db 61 TTTTCCCAACCGAGGCGAGGCGCTTCCGAGTCAATGCGCGGCTTCAATCTGCGGGTCTCT 120
Qy 310 ATGCACTGTAGAGGAGTCTGTGCTGTGAGGAGCGGCGCTGCGCGGCGGAGAGAGCT 369
Db 121 ATGCACTGTAGAGGAGTCTGTGCTGTGAGGAGCGGCGCTGCGCGGCGGAGAGAGCT 180
Qy 370 TTTGAGATTGCTCTGCGCAAGAAAGCAATTAAGTTTCTATTAAGCTGGGGAACCCCG 429
Db 181 TTTGAGATTGCTCTGCGCAAGAAAGCAATTAAGTTTCTATTAAGCTGGGGAACCCCG 240
Qy 430 ACTCTGCTGGCAAAACCCGTGTACATCGTCACTTTCTAAAGAATATAACATGTTGTCC 489
Db 241 ACTCTGCTGGCAAAACCCGTGTACATCGTCACTTTCTAAAGAATATAACATGTTGTCC 300
Qy 490 ATCAAGTCTGGAGAAAGAAATAGTCTTTACCTTTAGCTGCGAGAGTCTTGAAATCACTTT 549
Db 301 ATCAAGTCTGGAGAAAGAAATAGTCTTTACCTTTAGCTGCGAGAGTCTTGAAATCACTTT 360
Qy 550 GTCATATAGATCCAGAAAAATATATGACTGTATGTCAAGGCCCATGTCTTTGGGAGGTT 609
Db 361 GTCATATAGATCCAGAAAAATATATGACTGTATGTCAAGGCCCATGTCTTTGGGAGGTT 420
Qy 610 CAGCTTCAGCCCTCGACATCGTTGTTGGCTTACCCCTCAACAGAACTTTCACTGGAGATGC 669
Db 421 CAGCTTCAGCCCTCGACATCGTTGTTGGCTTACCCCTCAACAGAACTTTCACTGGAGATGC 480
Qy 670 AAAGCTATAGAGCATCGGTTTAAAGCTGAGCTGAGTTTCCATCCCTGCGCTGAGGAGATC 729
Db 481 AAAGCTATAGAGCATCGGTTTAAAGCTGAGCTGAGTTTCCATCCCTGCGCTGAGGAGATC 540
Qy 730 GGTCCGGGTGAGAGCTGCCCAAGCGAGTCACTCACTCATCAAGCGCGCAATCGATGCC 789
Db 541 GGTCCGGGTGAGAGCTGCCCAAGCGAGTCACTCACTCATCAAGCGCGCAATCGATGCC 600
Qy 790 ACCGTGTGAGAGTGGAACTTTCTGACAGCAATGGACATGTGTCCCGGATCAAGATGCA 849
Db 601 ACCGTGTGAGAGTGGAACTTTCTGACAGCAATGGACATGTGTCCCGGATCAAGATGCA 660
```

QY 850 GAAGAGTGAATGCGCTTACACCTTCCAGTGTCCACCCGAAATGTCCTCGGCTTC 909  
DB 661 GAAGAGTGAATGCGCTTACACCTTCCAGTGTCCACCCGAAATGTCCTCGGCTTC 720  
QY 910 AGGATTGCAAAACCGCTCATATATAAACGTCGTGACATCATGATGTCGTTGAGGGT 969  
DB 721 AGCATTTGCAAAACCGCTCATATATAAACGTCGTGACATCATGATGTCGTTGAGGGT 780  
QY 970 GAAGGCTCAGCAACCCGTGATGTCTGCACTACCAAGAGCTTCCCTGAGATGAGCTC 1029  
DB 781 GAAGGCTCAGCAACCCGTGATGTCTGCACTACCAAGAGCTTCCCTGAGATGAGCTC 840  
QY 1030 ATGACGTGGAGTTTGTCTCTCTGCAACCTTGCGGGCCAGGCTTCTCTCTCAACTTC 1089  
DB 841 ATACGTGGAGTTTGTCTCTCTGCAACCTTGCGGGCCAGGCTTCTCTCTCAACTTC 900  
QY 1090 AACCTTCCCACTGTAG 1149  
DB 901 AACCTTCCCACTGTAG 960  
QY 1150 ACCAACCAGAGGTGTTCAAGCTGAGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1209  
DB 961 ACCAACCAGAGGTGTTCAAGCTGAGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020  
QY 1210 AACCTTCTCTGCAAGGCTGTGACCAAGATGCCAAAGTCCAGGAGTCTTCCGCTGAG 1269  
DB 1021 AACCTTCTCTGCAAGGCTGTGACCAAGATGCCAAAGTCCAGGAGTCTTCCGCTGAG 1080  
QY 1270 TTCCAGTTTGTGTCACATCCCAAAATGAAGA 1306  
DB 1081 TTCCAGTTTGTGTCACATCCCAAAATGAAGA 1117

## RESULT 8

US-10-140-021-161

Sequence 161, Application US/10140021  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: DeForge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3330R1C167  
CURRENT APPLICATION NUMBER: US/10/140, 021  
CURRENT FILING DATE: 2002-05-06  
Prior Application removed - See Palm or File Wrapper  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 161  
LENGTH: 1376  
TYPE: DNA  
ORGANISM: Homo Sapien  
US-10-140-021-161

Query Match 18.1%; Score 1117; DB 13; Length 1376;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 190 CCAAGGCGAGAGCGACAGCTGCGCGCGGCTTTGGGCGCTTGAGGAGCGCGCGCTCCCAACCTGTG 249  
DB 1 CCAAGGCGAGAGCGAGCTGCGCGCGGCTTTGGGCGCTTGAGGAGCGCGCGCTCCCAACCTGTG 60  
QY 250 TTTTCCCCCAGAGGCGAGGCGCTCCCGGAGTCATGGCGGCTGAACCTGCGGGGCTCT 309  
DB 61 TTTTCCCCCAGAGGCGAGGCGCTCCCGGAGTCATGGCGGCTGAACCTGCGGGGCTCT 120  
QY 310 ATCGACATGCTGAGGGGTTCTGCTGCTGAGGCGGCGGCTGCGCGCGGCGAGAGCT 369  
DB 121 ATCGACATGCTGAGGGGTTCTGCTGCTGAGGCGGCGGCTGCGCGCGGCGAGAGCT 180  
QY 370 TTTGAGATTGCTCTGCGACAGAGAAACCAATTACAGTTCTCATTAAGCTGGAGACCCG 429  
DB 181 TTTGAGATTGCTCTGCGACAGAGAAACCAATTACAGTTCTCATTAAGCTGGAGACCCG 240  
QY 430 ACTGTGCTGGCAAAACCGCTGTACATCGCATTTCTTAAAGCATATACATGTTGTGTC 489  
DB 241 ACTGTGCTGGCAAAACCGCTGTACATCGCATTTCTTAAAGCATATACATGTTGTGTC 300  
QY 490 ATCAAGCTGAGAAAGAAATAGTCTTTACCTTTAGCTGCGAGAGTCTGAGATCACTT 549  
DB 301 ATCAAGCTGAGAAAGAAATAGTCTTTAGCTTACTGCGAGAGTCTGAGATCACTT 360  
QY 550 GTCATAGAGATCCAGAAATATTTAGCTGATGTCAAGGCCATGTCTTTGGAGGTT 609  
DB 361 GTCATAGAGATCCAGAAATATTTAGCTGATGTCAAGGCCATGTCTTTGGAGGTT 420  
QY 610 GAGCTTACGCGCTGACATCGTGTGCTGACCTTACCTTCAAGAGACCTTCACTGAGAGTGC 669  
DB 421 GAGCTTACGCGCTGACATCGTGTGCTGACCTTACCTTCAAGAGACCTTCACTGAGAGTGC 480  
QY 670 AAAGCTAATAGAGCATGCGTTTAAAGCTGAGCTGAGTTTCCATCCCTGCGTGAAGAGATC 729  
DB 481 AAAGCTAATAGAGCATGCGTTTAAAGCTGAGCTGAGTTTCCATCCCTGCGTGAAGAGATC 540  
QY 730 GGTCCGGGTGAGAGCTGCCAGAGAGTCACTCACTTCATCAGGCGCGGAATGATGCC 789  
DB 541 GGTCCGGGTGAGAGCTGCCAGAGAGTCACTCACTTCATCAGGCGCGGAATGATGCC 600  
QY 790 ACCGTGTGAGATGGAACCTTCTGCAAGATGGCACTGTCTCCCGATCAAGATGCA 849  
DB 601 ACCGTGTGAGATGGAACCTTCTGCAAGATGGCACTGTCTCCCGATCAAGATGCA 660  
QY 850 GAAGAGTGAATAATGCGCTTACACCTCCCATGTTCCACCAGAAATGTCTCCGGCTTC 909  
DB 661 GAAGAGTGAATAATGCGCTTACACCTCCCATGTTCCACCAGAAATGTCTCCGGCTTC 720  
QY 910 AGCATTTGCAAAACCGCTCATATATAAACGTCGTGACATCATGATGTCGTTGAGGGT 969  
DB 721 AGCATTTGCAAAACCGCTCATATATAAACGTCGTGACATCATGATGTCGTTGAGGGT 780  
QY 970 GAAGGCTCAGCAACCCGTGATGTCTGCACTACCAAGAGCTTCCCTGAGATGAGCTC 1029  
DB 781 GAAGGCTCAGCAACCCGTGATGTCTGCACTACCAAGAGCTTCCCTGAGATGAGCTC 840  
QY 1030 ATGAGTGGAGTTTGTCTCTCTGCAACCTTGCGGGCCAGGCTTCTCTCAACTTC 1089  
DB 841 ATGAGTGGAGTTTGTCTCTCTGCAACCTTGCGGGCCAGGCTTCTCTCAACTTC 900  
QY 1090 AACCTTCCCACTGTAG 1149  
DB 901 AACCTTCCCACTGTAG 960  
QY 1150 ACCAACCAGAGGTGTTCAAGCTGAGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1209  
DB 961 ACCAACCAGAGGTGTTCAAGCTGAGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020  
QY 1210 AACCTTCTCTGCAAGGCTGTGACCAAGATGCCAAAGTCCAGGAGTCTTCCGCTGAG 1269  
DB 1021 AACCTTCTCTGCAAGGCTGTGACCAAGATGCCAAAGTCCAGGAGTCTTCCGCTGAG 1080  
QY 1270 TTCCAGTTTGTGTCACATCCCAAAATGAAGA 1306



Db 1081 TTCCAAGTTTGGTCCCAATCATCCACAAATGAAGCA 1117

RESULT 9  
US-10-140-274-161

; Sequence 161, Application US/10140274  
; Publication No. US20030143674A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: Deforge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerltzen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330R1C161  
; CURRENT FILING DATE: 2002-05-06  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 161  
; LENGTH: 1376  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-140-274-161

Query Match 18.1%; Score 1117; DB 13; Length 1376;

Best Local Similarity 100.0%; Fred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 1117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 190 CCAGGCGGAGGCGAGCTGCGCGGCTTGGGCGCTGCGGCGCGCGCTCCCGACGCTG 249  
1 CCAGGCGGAGGCGAGCTGCGCGGCTTGGGCGCTGCGGCGCGCGCTCCCGACGCTG 60  
250 TTTTCCCAACCGAGGCGCGGCTTGGGCGCTGCGGCGCGCGCTCCCGACGCTGCT 309  
61 TTTTCCCAACCGAGGCGCGGCTTGGGCGCTGCGGCGCGCGCTCCCGACGCTGCT 120  
310 ATCGCACTGCTAGGGGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 369  
121 ATCGCACTGCTAGGGGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180  
370 TTTGAGATTGCTCTGCGCAAGAGCAATTAAGTCTGATAAGCTGGGGAACCCG 429  
181 TTTGAGATTGCTCTGCGCAAGAGCAATTAAGTCTGATAAGCTGGGGAACCCG 240  
430 ACTGCTGCGCAAAACCCCTGTTACATCGTCAATTTTAAAGACATATAACATGTTGTC 489  
241 ACTGCTGCGCAAAACCCCTGTTACATCGTCAATTTTAAAGACATATAACATGTTGTC 300  
480 ATCAAGTCTGGAAGAAATAGTCTTTTACCTTTAGCTGCAAGTCTGGAATCACTTT 549  
301 ATCAAGTCTGGAAGAAATAGTCTTTTACCTTTAGCTGCAAGTCTGGAATCACTTT 360  
550 GTCATAGAGATCCAGAAATATGTAAGTCTATGTCAGGCGCATGCTCTTTGGGAGGTT 609  
361 GTCATAGAGATCCAGAAATATGTAAGTCTATGTCAGGCGCATGCTCTTTGGGAGGTT 420  
610 CAGCTTCAAGCCCTCGACATCGTTGCTTACCTTCAACAGAACTTTCATCTGGAGTTC 669

Db 421 CAGCTTCAAGCCCTCGACATCGTTGCTTACCTCAACAGAACTTTCATCTGGAGTTC 480  
Qy 670 AAAGCTAATAAGAGATCGGTTTAAAGCTGAGCTGAGTTTTCATCCCTGCGCTGAGGAGATC 729  
Db 481 AAAGCTAATAAGAGATCGGTTTAAAGCTGAGCTGAGTTTTCATCCCTGCGCTGAGGAGATC 540  
Qy 730 GGTCCGGGTGAGAGCTGCGCAAGAGCTCACTCACTCCATCAGGCGCGAATGATGTC 789  
Db 541 GGTCCGGGTGAGAGCTGCGCAAGAGCTCACTCACTCCATCAGGCGCGAATGATGTC 600  
Qy 790 ACCGTGTCAGATGGAACCTTCTGCAAGCAATGCACTGTCTCCCGATCAAGATGCA 849  
Db 601 ACCGTGTCAGATGGAACCTTCTGCAAGCAATGCACTGTCTCCCGATCAAGATGCA 660  
Qy 850 GAAGAGTGAATAAGCCCTTACCTCCATGTTTCCACCCAGAAATGTCCTCGGCTTC 909  
Db 661 GAAGAGTGAATAAGCCCTTACCTCCATGTTTCCACCCAGAAATGTCCTCGGCTTC 720  
Qy 910 AGCATTGCAAAACCGCTCACTTATAAAAGCTCTGTCATCATGAGTCTGTGTTGAGGT 969  
Db 721 AGCATTGCAAAACCGCTCACTTATAAAAGCTCTGTCATCATGAGTCTGTGTTGAGGT 780  
Qy 970 GAAGCTCAAGCAACCTGATGCTGCACTCACTCAAGAGGCTTCCCTGAGATGAGTTC 1029  
Db 781 GAAGCTCAAGCAACCTGATGCTGCACTCACTCAAGAGGCTTCCCTGAGATGAGTTC 840  
Qy 1030 ATGAGTGGAGTGTGCTGCTGTCGACACCTGCGGCGAGGCTCTCTTCCCACTTC 1089  
Db 841 ATGAGTGGAGTGTGCTGCTGTCGACACCTGCGGCGAGGCTCTCTTCCCACTTC 900  
Qy 1090 AACCTCTCAACTGTGAGAGAGAGAGAGGCGGTTGAATACTACATCCCGGCTCCACC 1149  
Db 901 AACCTCTCAACTGTGAGAGAGAGAGAGGCGGTTGAATACTACATCCCGGCTCCACC 960  
Qy 1150 ACCAACCCTGAGGTTTCAAGCTGAGAGCAAGCAAGCTTGGGAATGCGGGGAATTC 1209  
Db 961 ACCAACCCTGAGGTTTCAAGCTGAGAGCAAGCAAGCTTGGGAATGCGGGGAATTC 1020  
Qy 1210 AACCTCTCTGCAAGGCTGAGCAAGATGCGCAAGTCCAGGAGTCTCCGGCTGAG 1269  
Db 1021 AACCTCTCTGCAAGGCTGAGCAAGATGCGCAAGTCCAGGAGTCTCCGGCTGAG 1080  
Qy 1270 TTCCAAGTTTGGTCCCAATCATCCCAAAATGAAGCA 1306  
Db 1081 TTCCAAGTTTGGTCCCAATCATCCCAAAATGAAGCA 1117

RESULT 10  
US-10-140-471-161

; Sequence 161, Application US/10140471  
; Publication No. US20030138887A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: Deforge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerltzen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330R1C163  
; CURRENT APPLICATION NUMBER: US/10/140,471

; CURRENT FILING DATE: 2002-05-06  
 ; Prior Application removed - See File Wrapper or Palm  
 ; NUMBER OF SEQ ID NOS: 550  
 ; SEQ ID NO 161  
 ; LENGTH: 1376  
 ; TYPE: DNA  
 ; ORGANISM: Homo Sapien  
 ; US-10-140-471-161

Query Match 18.1%; Score 1117; DB 13; Length 1376;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 190 CCAGGCGGAGGCGGAGGCGGCTCCGAGTCATGCGGCGCTGAACCTGCGGCTCT 249  
 Db 1 CCAGGCGGAGGCGGAGGCGGCTCCGAGTCATGCGGCGCTGAACCTGCGGCTCT 60  
 QY 250 TTTTCCCGGAGGCGGAGGCGGCTCCGAGTCATGCGGCGCTGAACCTGCGGCTCT 309  
 Db 61 TTTTCCCGGAGGCGGAGGCGGCTCCGAGTCATGCGGCGCTGAACCTGCGGCTCT 120  
 QY 310 ATGCACTGTAAGGGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 369  
 Db 121 ATGCACTGTAAGGGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180  
 QY 370 TTTGAGATTGCTCTGCGCAGAGAAACATTTACAGTTCTCATAAAGCTGGGACCCG 429  
 Db 181 TTTGAGATTGCTCTGCGCAGAGAAACATTTACAGTTCTCATAAAGCTGGGACCCG 240  
 QY 430 ACTCTGCTGGCAAAACCTGTTACATGTCATTTCAAAAGACATTAACATGTTGTC 489  
 Db 241 ACTCTGCTGGCAAAACCTGTTACATGTCATTTCAAAAGACATTAACATGTTGTC 300  
 QY 490 ATCAAGTCTGAGAAAGAAATAGTCTTTACCTTAGTGCAGAGTCCGAGATCACTT 549  
 Db 301 ATCAAGTCTGAGAAAGAAATAGTCTTTACCTTAGTGCAGAGTCCGAGATCACTT 360  
 QY 550 GTCATAGAGATCCAGAAATATTTGACTGTAATGTCAGGCCATGCTCTTTGGGAGGTT 609  
 Db 361 GTCATAGAGATCCAGAAATATTTGACTGTAATGTCAGGCCATGCTCTTTGGGAGGTT 420  
 QY 610 CAGCTTCAGCCCTGCAATCGTTGTTGCCATCCCTCAAGCAATCTTCACTGCGGATGTC 669  
 Db 421 CAGCTTCAGCCCTGCAATCGTTGTTGCCATCCCTCAAGCAATCTTCACTGCGGATGTC 480  
 QY 670 AAAGCTCATTAAGAGATCGGTTTAGAGCTGCAAGTTTCCATCCCTGCGCTGAGGAGATC 729  
 Db 481 AAAGCTCATTAAGAGATCGGTTTAGAGCTGCAAGTTTCCATCCCTGCGCTGAGGAGATC 540  
 QY 730 GGTCCGGGTGAGAGCTGCGCCAGAGATCACTCACTCACTCACTCACTCACTCACTCACT 789  
 Db 541 GGTCCGGGTGAGAGCTGCGCCAGAGATCACTCACTCACTCACTCACTCACTCACTCACT 600  
 QY 790 ACCGTGTGAGATCGGAACCTTCTGCAAGATGCACTGTCCTGCGGATCAAGATGCA 849  
 Db 601 ACCGTGTGAGATCGGAACCTTCTGCAAGATGCACTGTCCTGCGGATCAAGATGCA 660  
 QY 850 GAAGAGTGAATGAGCTTACACTCCATGCTTCAACCCAGAAATGCTCCGGCTTC 909  
 Db 661 GAAGAGTGAATGAGCTTACACTCCATGCTTCAACCCAGAAATGCTCCGGCTTC 720  
 QY 910 AGCATTTGCAAAACCGCTCATCTATTAACAGTCTGTCATCATGAGTCTGTTGAGGGT 969  
 Db 721 AGCATTTGCAAAACCGCTCATCTATTAACAGTCTGTCATCATGAGTCTGTTGAGGGT 780  
 QY 970 GAAGGCTCAGCAACCTGATGTCGCAATACCAAGAGCTTCCCTGAGATGAGCTC 1029  
 Db 781 GAAGGCTCAGCAACCTGATGTCGCAATACCAAGAGCTTCCCTGAGATGAGCTC 840  
 QY 1030 ATGACGTGAGATTTGCTGCTGCAACCTGCGGCGAGGCTCTCTTCTCAACTTC 1089  
 Db 841 ATGACGTGAGATTTGCTGCTGCAACCTGCGGCGAGGCTCTCTTCTCAACTTC 900

QY 1090 AACCTCTCCAACTGTGAGAGAGAGAGAGGCGGTTGAATATCTACATCCCGGCTTCACC 1149  
 Db 901 AACCTCTCCAACTGTGAGAGAGAGAGAGGCGGTTGAATATCTACATCCCGGCTTCACC 960  
 QY 1150 ACCAAGCCGAGAGTGTTCAGCTGAGAGAGAGAGAGGCGGTTGAATATCTACATCCCGGCTTCACC 1209  
 Db 961 ACCAAGCCGAGAGTGTTCAGCTGAGAGAGAGAGAGGCGGTTGAATATCTACATCCCGGCTTCACC 1020  
 QY 1210 AACCTCTCTGCAAGGCTGTGAGAGAGAGAGAGGCGGTTGAATATCTACATCCCGGCTTCACC 1269  
 Db 1021 AACCTCTCTGCAAGGCTGTGAGAGAGAGAGAGGCGGTTGAATATCTACATCCCGGCTTCACC 1080  
 QY 1270 TTTGAGATTGCTCTGCGCAGAGAAACATTTACAGTTCTCATAAAGCTGGGACCCG 1306  
 Db 1081 TTTGAGATTGCTCTGCGCAGAGAAACATTTACAGTTCTCATAAAGCTGGGACCCG 1117

RESULT 11

US-10-140-807-161  
 ; Sequence 161, Application US/10140807  
 ; Publication No. US20030134354A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Baker, Kevin P.  
 ; APPLICANT: Beresini, Maureen  
 ; APPLICANT: DeForge, Laura  
 ; APPLICANT: Desnoyers, Luc  
 ; APPLICANT: Filvaroff, Ellen  
 ; APPLICANT: Gao, Wei-Qiang  
 ; APPLICANT: Gerlitsen, Mary E.  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Godowski, Paul J.  
 ; APPLICANT: Gurney, Austin L.  
 ; APPLICANT: Sherwood, Steven  
 ; APPLICANT: Smith, Victoria  
 ; APPLICANT: Stewart, Timothy A.  
 ; APPLICANT: Tumas, Daniel  
 ; APPLICANT: Watanabe, Colin K.  
 ; APPLICANT: Wood, William  
 ; APPLICANT: Zhang, Zemin  
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
 ; FILE REFERENCE: P3330R1C174  
 ; CURRENT APPLICATION NUMBER: US/10/140,807  
 ; CURRENT FILING DATE: 2002-05-07  
 ; Prior Application removed - See File Wrapper or Palm  
 ; NUMBER OF SEQ ID NOS: 550  
 ; SEQ ID NO 161  
 ; LENGTH: 1376  
 ; TYPE: DNA  
 ; ORGANISM: Homo Sapien  
 ; US-10-140-807-161

Query Match 18.1%; Score 1117; DB 13; Length 1376;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 190 CCAGGCGGAGGCGGAGGCGGCTCCGAGTCATGCGGCGCTGAACCTGCGGCTCT 249  
 Db 1 CCAGGCGGAGGCGGAGGCGGCTCCGAGTCATGCGGCGCTGAACCTGCGGCTCT 60  
 QY 250 TTTTCCCGGAGGCGGAGGCGGCTCCGAGTCATGCGGCGCTGAACCTGCGGCTCT 309  
 Db 61 TTTTCCCGGAGGCGGAGGCGGCTCCGAGTCATGCGGCGCTGAACCTGCGGCTCT 120  
 QY 310 ATGCACTGTAAGGGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 369  
 Db 121 ATGCACTGTAAGGGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180  
 QY 370 TTTGAGATTGCTCTGCGCAGAGAAACATTTACAGTTCTCATAAAGCTGGGACCCG 429  
 Db 181 TTTGAGATTGCTCTGCGCAGAGAAACATTTACAGTTCTCATAAAGCTGGGACCCG 240  
 QY 430 ACTCTGCTGGCAAAACCTGTTACATGTCATTTCAAAAGACATTAACATGTTGTC 489

Db 241 ACTCTGCTGGCAAAACCCCTGTTACATCGTCTTAAAGACATATAACATGTTGTC 300  
Qy 490 ATCAAGTCTGGAGAAAGATTAATCTTTACTTTAGTGTGCAAGTCCGAGATCACTT 549  
Db 301 ATCAAGTCTGGAGAAAGATTAATCTTTACTTTAGTGTGCAAGTCCGAGATCACTT 360  
Qy 550 GTCATAGAGATCCAGAAATAATTAATGATGATGTCAGGCCCATGTCCTTTGGAGAGTT 609  
Db 361 GTCATAGAGATCCAGAAATAATTAATGATGATGTCAGGCCCATGTCCTTTGGAGAGTT 420  
Qy 610 CAGCTTCAGCCCTCGACATGTTGTTGCTTACCTTCACAGAACTTTCACTGGAGATGTC 669  
Db 421 CAGCTTCAGCCCTCGACATGTTGTTGCTTACCTTCACAGAACTTTCACTGGAGATGTC 480  
Qy 670 AAAGCTCATAGAGATCCAGAAATAATTAATGATGATGTCAGGCCCATGTCCTTTGGAGAGTT 729  
Db 481 AAAGCTCATAGAGATCCAGAAATAATTAATGATGATGTCAGGCCCATGTCCTTTGGAGAGTT 540  
Qy 730 GGTCCGGGTGAGAGTGTCCAGACGAGATCACTCACTCATGAGCGGCGGAATGATGCC 789  
Db 541 GGTCCGGGTGAGAGTGTCCAGACGAGATCACTCACTCATGAGCGGCGGAATGATGCC 600  
Qy 790 ACCGTGTGAGAGATCCAGAACTTTGTCAGCAATGGCACTGTGTCCGGATCAAGATGCA 849  
Db 601 ACCGTGTGAGAGATCCAGAACTTTGTCAGCAATGGCACTGTGTCCGGATCAAGATGCA 660  
Qy 850 GAAGAGTGAATAATGAGCTTTACACTCCCATGTTCCACCCAGAAATGTCTCCGGCTTC 909  
Db 661 GAAGAGTGAATAATGAGCTTTACACTCCCATGTTCCACCCAGAAATGTCTCCGGCTTC 720  
Qy 910 AGCATTTGCAAAACCGCTCATCTATAAAAGTCTGTGATGATGAGATGTTTGAAGGT 969  
Db 721 AGCATTTGCAAAACCGCTCATCTATAAAAGTCTGTGATGATGAGATGTTTGAAGGT 780  
Qy 970 GAAGGCTCAGCAACCTGATGTCCTGCACTACCCAGAAAGCTTCCTGAGATGAGCTC 1029  
Db 781 GAAGGCTCAGCAACCTGATGTCCTGCACTACCCAGAAAGCTTCCTGAGATGAGCTC 840  
Qy 1030 ATGAGTGTGAGATGTTGTCCTGTCACACTGTCGGGCGAGGCTCTCTCTCAACTTC 1089  
Db 841 ATGAGTGTGAGATGTTGTCCTGTCACACTGTCGGGCGAGGCTCTCTCTCAACTTC 900  
Qy 1090 AACCTCTCCAACTGTGAGAGAGAGAGAGGAGGTTGAATCTACATCCCGGCTCCACC 1149  
Db 901 AACCTCTCCAACTGTGAGAGAGAGAGAGGAGGTTGAATCTACATCCCGGCTCCACC 960  
Qy 1150 ACCAACCCCGAGGTTTCAAGCTGAGAGCAAGCAAGCTGAGAACATGAGCGGGAATTC 1209  
Db 961 ACCAACCCCGAGGTTTCAAGCTGAGAGCAAGCAAGCTGAGAACATGAGCGGGAATTC 1020  
Qy 1210 AACCTCTCTCTGCAAGAGCTGTGACCAAGATGTCGAGGATCTCTCGGCTGCAG 1269  
Db 1021 AACCTCTCTCTGCAAGAGCTGTGACCAAGATGTCGAGGATCTCTCGGCTGCAG 1080  
Qy 1270 TTCCAAGTTTGGTCCAAATCCAAATAATGAAGCA 1306  
Db 1081 TTCCAAGTTTGGTCCAAATCCAAATAATGAAGCA 1117

RESULT 12  
US-10-140-922-161

Sequence 161, Application US/10140922  
Publication No. US20030138889A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: DeForge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvarolf, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Geritsen, Mary E.  
APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3330R1C179  
CURRENT APPLICATION NUMBER: US/10/140,922  
CURRENT FILING DATE: 2002-05-07  
Prior Application removed - See Palm or File Wrapper  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 161  
LENGTH: 1376  
TYPE: DNA  
ORGANISM: Homo Sapien  
US-10-140-922-161

Query Match 18.1%; Score 1117; DB 13; Length 1376;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 190 CCAAGGCGGAGAGCGAGCTGCGCGGCTTTGGGCGCTGCGGCGCGGCTTCCCAACCTG 249  
Db 1 CCAAGGCGGAGAGCGAGCTGCGCGGCTTTGGGCGCTGCGGCGCGGCTTCCCAACCTG 60  
Qy 250 TTTTCCCAACCGAGCGAGGCGTCCGAGATCATGCGCGGCTGAACTGCGGAGCTCT 309  
Db 61 TTTTCCCAACCGAGCGAGGCGTCCGAGATCATGCGCGGCTGAACTGCGGAGCTCT 120  
Qy 310 ATGCACTGTAGAGGTTCTGCTGCTGAGGCGGCGGCTGCGGCGGCGGAGCAACT 369  
Db 121 ATGCACTGTAGAGGTTCTGCTGCTGAGGCGGCGGCTGCGGCGGCGGAGCAACT 180  
Qy 370 TTTGAGATTTGCTTCCCAACGAGAAAGCAATTAAGTTTCAATAAAGCTGGGAGCCCCG 429  
Db 181 TTTGAGATTTGCTTCCCAACGAGAAAGCAATTAAGTTTCAATAAAGCTGGGAGCCCCG 240  
Qy 430 ACTGTGCTGGCAAAACCTGTTTACATGCTCAATTTCTAAAGACATATAACATGTTGTC 489  
Db 241 ACTGTGCTGGCAAAACCTGTTTACATGCTCAATTTCTAAAGACATATAACATGTTGTC 300  
Qy 490 ATCAAGTCTGAGAAAGATGATCTTTTACTTTAGCTGCGAGATCTGAGAAATCACTT 549  
Db 301 ATCAAGTCTGAGAAAGATGATGTTTACTTTTACTTTAGCTGCGAGATCTGAGAAATCACTT 360  
Qy 550 GTCATAGAGATCCAGAAATAATTAATGATGATGTCAGGCCCATGTCCTTTGGAGAGTT 609  
Db 361 GTCATAGAGATCCAGAAATAATTAATGATGATGTCAGGCCCATGTCCTTTGGAGAGTT 420  
Qy 610 CAGCTTCAGCCCTCGACATGTTGTTGCTTACCTTCACAGAACTTTCACTGGAGATGTC 669  
Db 421 CAGCTTCAGCCCTCGACATGTTGTTGCTTACCTTCACAGAACTTTCACTGGAGATGTC 480  
Qy 670 AAAGCTCATAGAGATCCAGAAATAATTAATGATGATGTCAGGCCCATGTCCTTTGGAGAGTT 729  
Db 481 AAAGCTCATAGAGATCCAGAAATAATTAATGATGATGTCAGGCCCATGTCCTTTGGAGAGTT 540  
Qy 730 GGTCCGGGTGAGAGTGTCCAGACGAGATCACTCACTCATGAGCGGCGGAATGATGCC 789  
Db 541 GGTCCGGGTGAGAGTGTCCAGACGAGATCACTCACTCATGAGCGGCGGAATGATGCC 600  
Qy 790 ACCGTGTGAGAGATCCAGAACTTTGTCAGCAATGGCACTGTGTCCGGATCAAGATGCA 849  
Db 601 ACCGTGTGAGAGATCCAGAACTTTGTCAGCAATGGCACTGTGTCCGGATCAAGATGCA 660  
Qy 850 GAAGAGTGAATAATGAGCTTTACACTCCCATGTTCCACCCAGAAATGTCTCCGGCTTC 909  
Db 661 GAAGAGTGAATAATGAGCTTTACACTCCCATGTTCCACCCAGAAATGTCTCCGGCTTC 720

QY 910 AGCATTTGCAAAACCGCTCATCTATATAAAAGTCTGTGCATCATGAGTCTGTGTTGAGGCT 969  
DB 721 AGCATTTGCAAAACCGCTCATCTATATAAAAGTCTGTGCATCATGAGTCTGTGTTGAGGCT 780  
QY 970 GAAGGCTGCAAAACCGCTCATCTATATAAAAGTCTGTGCATCATGAGTCTGTGTTGAGGCT 1029  
DB 781 GAAGGCTGCAAAACCGCTCATCTATATAAAAGTCTGTGCATCATGAGTCTGTGTTGAGGCT 840  
QY 1030 ATGACGTGCAAAACCGCTCATCTATATAAAAGTCTGTGCATCATGAGTCTGTGTTGAGGCT 1089  
DB 841 ATGACGTGCAAAACCGCTCATCTATATAAAAGTCTGTGCATCATGAGTCTGTGTTGAGGCT 900  
QY 1090 AACCTCTCCAACTGTGAG 1149  
DB 901 AACCTCTCCAACTGTGAG 960  
QY 1150 ACCAACCCGAGGTGTTCAAGCTGAG 1209  
DB 961 ACCAACCCGAGGTGTTCAAGCTGAG 1020  
QY 1210 AACCTCTCTGCAAGAGCTGTGACCAAGATGCCCCAAGTCCAGAGGATCCCTCCGCTGCAG 1269  
DB 1021 AACCTCTCTGCAAGAGCTGTGACCAAGATGCCCCAAGTCCAGAGGATCCCTCCGCTGCAG 1080  
QY 1270 TTCCAAGTTTGGTCCCAACATCCCAAAATGAAAGCA 1306  
DB 1081 TTCCAAGTTTGGTCCCAACATCCCAAAATGAAAGCA 1117

RESULT 13  
US-10-140-924-161  
Sequence 161, Application US/10140924  
Publication No. US20030134355A1  
GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: DeForge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerltzen, Mary E.  
APPLICANT: Goddard, Audrey J.  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Thomas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3330R1C177  
CURRENT APPLICATION NUMBER: US/10/140, 924  
PRIORITY FILING DATE: 2002-05-07  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 161  
LENGTH: 1376  
TYPE: DNA  
ORGANISM: Homo Sapien  
US-10-140-924-161

Query Match 18.1%; Score 1117; DB 13; Length 1376;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 190 CCAGGCGGAGGAGCAGCTGCGCGGCGCTTGGGCGCTGGGCGCGCGCTCCCAACCGTGG 249  
DB 1 CCAGGCGGAGGAGCAGCTGCGCGGCGCTTGGGCGCTGGGCGCGCGCTCCCAACCGTGG 60

QY 250 TTTTCCCAACCGAGGCGGAGGCTCCCGAGTCAAGGCGGCTGAACTGCGGGGCTCT 309  
DB 61 TTTTCCCAACCGAGGCGGAGGCTCCCGAGTCAAGGCGGCTGAACTGCGGGGCTCT 120  
QY 310 ATGCGACTGCTGAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 369  
DB 121 ATGCGACTGCTGAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180  
QY 370 TTTGAGATTTGCTCTGCAAG 429  
DB 181 TTTGAGATTTGCTCTGCAAG 240  
QY 430 ACTCTGCTGCAAAACCGCTGTTACATGCTGATTTTCTAAAGACATTAATCAATGTTGCTC 489  
DB 241 ACTCTGCTGCAAAACCGCTGTTACATGCTGATTTTCTAAAGACATTAATCAATGTTGCTC 300  
QY 490 ATCAAGTCTGAG 549  
DB 301 ATCAAGTCTGAG 360  
QY 550 GTCATAGAGATCCAG 609  
DB 361 GTCATAGAGATCCAG 420  
QY 610 CAGCTTCAGGCTTCCAGATCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 669  
DB 421 CAGCTTCAGGCTTCCAGATCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
QY 670 AAAGCTCATTAAGAGATCGGTTTGAAGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 729  
DB 481 AAAGCTCATTAAGAGATCGGTTTGAAGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540  
QY 730 GGTTCGGGTGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 789  
DB 541 GGTTCGGGTGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
QY 790 ACCGTGTGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAG 849  
DB 601 ACCGTGTGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAG 660  
QY 850 GAAGAGTGAAG 909  
DB 661 GAAGAGTGAAG 720  
QY 910 AGCATTTGCAAAACCGCTCATCTATATAAAAGTCTGTGCATCATGAGTCTGTGTTGAGGCT 969  
DB 721 AGCATTTGCAAAACCGCTCATCTATATAAAAGTCTGTGCATCATGAGTCTGTGTTGAGGCT 780  
QY 970 GAAGGCTGCAAAACCGCTCATCTATATAAAAGTCTGTGCATCATGAGTCTGTGTTGAGGCT 1029  
DB 781 GAAGGCTGCAAAACCGCTCATCTATATAAAAGTCTGTGCATCATGAGTCTGTGTTGAGGCT 840  
QY 1030 ATGACGTGCAAAACCGCTCATCTATATAAAAGTCTGTGCATCATGAGTCTGTGTTGAGGCT 1089  
DB 841 ATGACGTGCAAAACCGCTCATCTATATAAAAGTCTGTGCATCATGAGTCTGTGTTGAGGCT 900  
QY 1090 AACCTCTCCAACTGTGAG 1149  
DB 901 AACCTCTCCAACTGTGAG 960  
QY 1150 ACCAACCCGAGGTGTTCAAGCTGAG 1209  
DB 961 ACCAACCCGAGGTGTTCAAGCTGAG 1020  
QY 1210 AACCTCTCTGCAAGAGCTGTGACCAAGATGCCCCAAGTCCAGAGGATCCCTCCGCTGCAG 1269  
DB 1021 AACCTCTCTGCAAGAGCTGTGACCAAGATGCCCCAAGTCCAGAGGATCCCTCCGCTGCAG 1080  
QY 1270 TTCCAAGTTTGGTCCCAACATCCCAAAATGAAAGCA 1306  
DB 1081 TTCCAAGTTTGGTCCCAACATCCCAAAATGAAAGCA 1117



LENGTH: 1376  
TYPE: DNA  
ORGANISM: Homo Sapien  
US-10-141-698-161

Query Match 18.1%; Score 1117; DB 13; Length 1376;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 1117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

190 CCAGGGCGAGCGCAGCTGCGCGGGCTTGGGCGCTGGGGCGCGCCCTCCACCCTG 249  
1 CCAGGGCGAGCGCAGCTGCGCGGGCTTGGGCGCTGGGGCGCGCCCTCCACCCTG 60

250 TTTTCCCAACCGAGCGCGGCTCCCGAGTCATGCGCGCTGACCTGCGGGCTCT 309  
61 TTTTCCCAACCGAGCGCGGCTCCCGAGTCATGCGCGCTGACCTGCGGGCTCT 120

310 ATCGCACTGCTAGGGGCTCTGCTGCTGGGGTGGGGCGCGCGCTGCGGGCGAGAGCT 369  
121 ATCGCACTGCTAGGGGCTCTGCTGCTGGGGTGGGGCGCGCGCTGCGGGCGAGAGCT 180

370 TTTGAGATTGCTCTGCGCAGAGAAAGCAATTACAGTTCTCATAAAGCTGGGGAACCCG 429  
181 TTTGAGATTGCTCTGCGCAGAGAAAGCAATTACAGTTCTCATAAAGCTGGGGAACCCG 240

430 ACTCTGCTGGCAAAACCTGTTACATGTCATTTCTAAAGACATTAACATGTTGCC 489  
241 ACTCTGCTGGCAAAACCTGTTACATGTCATTTCTAAAGACATTAACATGTTGCC 300

490 ATCAAGTCTGGAAGAAATAGCTTACCTTAAAGTGGAGTCTGAGAAATCACTT 549  
301 ATCAAGTCTGGAAGAAATAGCTTACCTTAAAGTGGAGTCTGAGAAATCACTT 360

550 GTCATAGAGATCCAGAAATAATGACTGATGTGAGGCCCATGCTTTTGGGAGATT 609  
361 GTCATAGAGATCCAGAAATAATGACTGATGTGAGGCCCATGCTTTTGGGAGATT 420

610 CAGCTTCAGCCCTCGACATGCTGTTGCTTACCTCAACAGAACTTTCATCTGGGATGTC 669  
421 CAGCTTCAGCCCTCGACATGCTGTTGCTTACCTCAACAGAACTTTCATCTGGGATGTC 480

670 AAAGCTCATTAAGAGATCGGTTAGAGCTGAGCTTTTCAATCCCTGCGTGGAGGAGATC 729  
481 AAAGCTCATTAAGAGATCGGTTAGAGCTGAGCTTTTCAATCCCTGCGTGGAGGAGATC 540

730 GGTCCGGGTGAGAGCTGCCAGAGAGTCACTCACTCATCAGCGCGGAAATCGATGCC 789  
541 GGTCCGGGTGAGAGCTGCCAGAGAGTCACTCACTCATCAGCGCGGAAATCGATGCC 600

790 ACCGTGTCAGAGATCGGAACCTTCTGAGCAATGAGCACTGTGTCCGGATCAAGATGCA 849  
601 ACCGTGTCAGAGATCGGAACCTTCTGAGCAATGAGCACTGTGTCCGGATCAAGATGCA 660

850 GAAAGGATGAAATGCGCTTACACCTCCAGTGTTCACCCAGAAATGTCTCCGCTTC 909  
661 GAAAGGATGAAATGCGCTTACACCTCCAGTGTTCACCCAGAAATGTCTCCGCTTC 720

910 AGCATTTGCAAAACCGCTCATCTAATAACGTCTGTGATCATGAGTCTGTGTTGAGGAT 969  
721 AGCATTTGCAAAACCGCTCATCTAATAACGTCTGTGATCATGAGTCTGTGTTGAGGAT 780

970 GAAAGCTCAGCAACCTGTATGTCTGCAACTACCCAGAAAGCTTCCCTGAGATGAGCTC 1029  
781 GAAAGCTCAGCAACCTGTATGTCTGCAACTACCCAGAAAGCTTCCCTGAGATGAGCTC 840

1030 ATGACCTGCGCAATTTGCTCTTCTGCAACTGCGGGCGAGGGTCTCTTCTCAACTTC 1089  
841 ATGACCTGCGCAATTTGCTCTTCTGCAACTGCGGGCGAGGGTCTCTTCTCAACTTC 900

1090 AACCTCTCAACTGTGAGAGAAAGAGAGCGGGTTGAATACTACATCCCGGGCTCCACC 1149  
901 AACCTCTCAACTGTGAGAGAAAGAGAGCGGGTTGAATACTACATCCCGGGCTCCACC 960

QY 1150 ACCAACCAGGCTGCTTCAAGTGAAGAGCAAGCAGCTTGGGAAATGCGGGGAACCTTC 1209  
DB 961 ACCAACCAGGCTGCTTCAAGTGAAGAGCAAGCAGCTTGGGAAATGCGGGGAACCTTC 1020

QY 1210 AACCTCTCTGCAAGGCTGTGACCAAGATGCCAAAGTCCAGGATCCTCCGGCTGCAG 1269  
DB 1021 AACCTCTCTGCAAGGCTGTGACCAAGATGCCAAAGTCCAGGATCCTCCGGCTGCAG 1080

QY 1270 TTCCAAGTTTGTGTCACATCCACAAATGAAAGCA 1306  
DB 1081 TTCCAAGTTTGTGTCACATCCACAAATGAAAGCA 1117

Search completed: February 20, 2004, 19:44:59  
Job time : 1986 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using SW model

Run on: February 20, 2004, 13:49:26 ; Search time 87 Seconds

(without alignments)  
1525.235 Million cell updates/sec

Title: US-09-899-569a-4

Perfect score: 836  
Sequence: 1 MAGLNCVSLALGVLLGA.....SSKQDIPLNTQBPMPBP 836

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 1107863 seqs, 158726573 residues

Word size: 0  
Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

A\_GeneSeq\_19Jun03:\*

1: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1980.DAT:\*  
2: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1981.DAT:\*  
3: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1982.DAT:\*  
4: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1983.DAT:\*  
5: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1984.DAT:\*  
6: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1985.DAT:\*  
7: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1986.DAT:\*  
8: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1987.DAT:\*  
9: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1988.DAT:\*  
10: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1989.DAT:\*  
11: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1990.DAT:\*  
12: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1991.DAT:\*  
13: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1992.DAT:\*  
14: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1993.DAT:\*  
15: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1994.DAT:\*  
16: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1995.DAT:\*  
17: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1996.DAT:\*  
18: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1997.DAT:\*  
19: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1998.DAT:\*  
20: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1999.DAT:\*  
21: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2000.DAT:\*  
22: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2001.DAT:\*  
23: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2002.DAT:\*  
24: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	836	100.0	836	23	ABP69553
2	836	100.0	836	23	ABP69553
3	836	100.0	836	23	ABP69553
4	836	100.0	836	23	ABP69553
5	836	100.0	836	23	ABP69553
6	836	100.0	836	23	ABP69553
7	836	100.0	836	23	ABP69553
8	836	100.0	836	23	ABP69553
9	836	100.0	836	23	ABP69553

10	211	25.2	709	21	AAV91456
11	159	19.0	159	21	AAV91593
12	143	17.1	143	22	ABG00296
13	121	14.5	121	22	AAV23969
14	74	8.9	116	22	AAV90342
15	39	4.7	167	22	AAV69483
16	9	1.1	19	21	AAV69781
17	9	1.1	75	24	ABP56972
18	9	1.1	215	24	ABP71325
19	9	1.1	215	24	ABP56994
20	9	1.1	293	24	ABP57488
21	9	1.1	2359	24	ABP56959
22	9	1.1	2360	24	ABP56972
23	9	1.1	2661	24	ABP56958
24	8	1.0	62	23	ABP53596
25	8	1.0	70	24	ABP56985
26	8	1.0	71	20	AAV30830
27	8	1.0	75	23	ABP10049
28	8	1.0	78	22	ABG56028
29	8	1.0	78	22	ABG40578
30	8	1.0	78	22	ABG24869
31	8	1.0	78	22	AAV61436
32	8	1.0	78	22	AAV74226
33	8	1.0	78	22	AAV34338
34	8	1.0	78	22	AAV44180
35	8	1.0	132	22	AAO00490
36	8	1.0	133	22	ABG22899
37	8	1.0	147	22	AAE03940
38	8	1.0	156	22	AAU18521
39	8	1.0	179	22	AAV83192
40	8	1.0	187	22	AAV98984
41	8	1.0	193	22	AAU48107
42	8	1.0	213	21	AAV18289
43	8	1.0	221	20	AAV88495
44	8	1.0	221	21	AAV25793
45	8	1.0	221	21	AAV32203

#### ALIGNMENTS

RESULT 1	ABP69553	standard; Protein; 836 AA.
ID	ABP69553	
AC	ABP69553	
DT	20-JAN-2003	(first entry)
DE	Human polypeptide SEQ ID NO 1600.	
XX	Human; genome mapping; gene therapy; food supplement; virus; fungus;	
XX	cell-proliferative disorder; neurodegenerative disease; bacterial;	
XX	Parkinson's disease; Alzheimer's disease; autoimmune disease;	
XX	multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;	
XX	arthritis; cytostatic; immunomodulator; nocotropic; neuroprotective;	
XX	antiparkinsonian; antidiabetic; immunosuppressive; dermatological;	
XX	haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;	
XX	antiarthritic.	
OS	Homo sapiens.	
XX	WO200270539-A2.	
XX	12-SEP-2002.	
XX	05-MAR-2002; 2002WO-US05095.	
XX	05-MAR-2001; 2001US-0799451.	
XX	(HYSE-) HYSEQ INC.	
XX	Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;	

PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;  
PI Wehman T, Wang J, Wang D, Drmanac RT;  
DR MPI: 2002-759812/82.  
XX N-PSDB; AB211770.  
XX  
XX New polynucleotides comprising sequences assembled from expressed  
PT sequence tags (ESTs), useful for treating cell-proliferative,  
PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or  
PT platelet or coagulation disorders -  
XX  
XX  
PS Claim 9; SEQ ID NO 1600; 1012pp + Sequence Listing; English.  
XX  
XX The invention relates to an isolated polynucleotide (1) comprising a  
CC nucleotide sequence selected from any of 948 sequences  
CC (AB21119-AB212066) or their mature protein coding portion, active domain  
CC coding protein or complementary sequences. The polynucleotides are useful  
CC for identifying expressed genes or for physical mapping of human genome.  
CC The encoded polypeptides (ABP68902-ABP69849) are useful as molecular  
CC weight markers, as a food supplement, for generating antibodies, in  
CC medical imaging, screening and diagnostic assays and for treating  
CC cell-proliferative disorders (cancer), neurodegenerative diseases  
CC (Parkinson's or Alzheimer's disease), autoimmune diseases (multiple  
CC sclerosis, diabetes, lupus) genetic disorders, myeloid or lymphoid  
CC disorders, platelet or coagulation disorders, wound, burns, incision,  
CC ulcers, liver or lung fibrosis, infections (bacterial, viral, fungal,  
CC parasitic), arthritis, etc.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX

Sequence 836 AA;  
Query Match 100.0%; Score 836; DB 23; Length 836;  
Best Local Similarity 100.0%; Pred No. 0;  
Matches 836; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGLNCVSIALLGVLLGAAALPRGAFAEIALPRESNITVLIKLGTPLAKPCYIV 60  
DB 1 MAGLNCVSIALLGVLLGAAALPRGAFAEIALPRESNITVLIKLGTPLAKPCYIV 60  
QY 61 SKRHITMLSIKSGERIVFTFSCSPENHVFYIEOKNIDCMGCPGPEVQLQPSLTLPT 120  
DB 61 SKRHITMLSIKSGERIVFTFSCSPENHVFYIEOKNIDCMGCPGPEVQLQPSLTLPT 120  
QY 121 LNRFTIWDVKAHKSIGLELOFSIPRLKQIGPESCPDGVTHSISGRIDATVAIGTFCN 180  
DB 121 LNRFTIWDVKAHKSIGLELOFSIPRLKQIGPESCPDGVTHSISGRIDATVAIGTFCN 180  
QY 181 GTVSRIKMOEGVMALHLPWFHPRNVSGFSIANRSSIKRLCTIESVFEBSGATLMSANY 240  
DB 181 GTVSRIKMOEGVMALHLPWFHPRNVSGFSIANRSSIKRLCTIESVFEBSGATLMSANY 240  
QY 241 PGCFPEDELMTWQFVPAHLRASVSFLNPNLSCGERKEVEYYIGSTTNPEVFLBOK 300  
DB 241 PGCFPEDELMTWQFVPAHLRASVSFLNPNLSCGERKEVEYYIGSTTNPEVFLBOK 300  
QY 241 PGCFPEDELMTWQFVPAHLRASVSFLNPNLSCGERKEVEYYIGSTTNPEVFLBOK 300  
DB 241 PGCFPEDELMTWQFVPAHLRASVSFLNPNLSCGERKEVEYYIGSTTNPEVFLBOK 300  
QY 301 QPENMAGNPNLSLOGCDQDAOSPGILRLOFVLVOPHNSNKIYVVDLSNERAMSLTIE 360  
DB 301 QPENMAGNPNLSLOGCDQDAOSPGILRLOFVLVOPHNSNKIYVVDLSNERAMSLTIE 360  
QY 361 PRVVKOSRKFPVPCFVCLSESRITCSNLTLSGSKHISFLCDLTLMMNVEKTIISCTH 420  
DB 361 PRVVKOSRKFPVPCFVCLSESRITCSNLTLSGSKHISFLCDLTLMMNVEKTIISCTH 420  
QY 421 RYCORSYSLOVPSDILHLPVELHDFPSKMLVPRKRLSLVLPVAKLQOHTHEKPCNTSF 480  
DB 421 RYCORSYSLOVPSDILHLPVELHDFPSKMLVPRKRLSLVLPVAKLQOHTHEKPCNTSF 480  
QY 481 SYLVAAIIPBODLYFSGFCPGSGSIKQIQVKNISVLTLTFTAPBFOBASQGLTVSFITY 540  
DB 481 SYLVAAIIPBODLYFSGFCPGSGSIKQIQVKNISVLTLTFTAPBFOBASQGLTVSFITY 540

QY 541 FKESGVFTVPTDTSKAVYLRIPNMDRGLPSLTUSWNISVPPDQVACLTFFKERSGVQC 600  
DB 541 FKESGVFTVPTDTSKAVYLRIPNMDRGLPSLTUSWNISVPPDQVACLTFFKERSGVQC 600  
QY 601 TGRAFMIIOEORTRAEELFSUDEVLPRPSFHHHSFWNINISCSPTSGKQDLTLSSVLT 660  
DB 601 TGRAFMIIOEORTRAEELFSUDEVLPRPSFHHHSFWNINISCSPTSGKQDLTLSSVLT 660  
QY 661 PRTVDLTVILAAVGGVLLSALGLIICVAKKKKKTKNGPAGVIGNININTEMPQPK 720  
DB 661 PRTVDLTVILAAVGGVLLSALGLIICVAKKKKKTKNGPAGVIGNININTEMPQPK 720  
QY 721 KFOGRKNDSHVAVIEDTWYGHLLQDSSGSPLOPEVDYTRPQGTMGVCPSPSPITC 780  
DB 721 KFOGRKNDSHVAVIEDTWYGHLLQDSSGSPLOPEVDYTRPQGTMGVCPSPSPITC 780  
QY 781 SRAPIAKLATEEPPRSPSESEPEYTSHPNNGVSSKDNDIPLINQEPMEPAE 836  
DB 781 SRAPIAKLATEEPPRSPSESEPEYTSHPNNGVSSKDNDIPLINQEPMEPAE 836

RESULT 2  
AAM49641  
ID AAM49641 standard; Protein; 836 AA.  
XX  
XX AAM49641;  
AC  
XX  
XX 17-MAY-2002 (first entry)  
DT  
XX  
XX Human tumour-associated antigen B345 protein SEQ ID NO 4.  
DE  
XX  
XX

XX Tumour-associated antigen; human; B345; cytostatic; cell communication;  
KW cell interaction; signal transduction; metastasis; cancer; colon;  
KW immunotherapy; carcinoma; lung; diagnosis.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO200204508-A1.  
PN  
XX  
XX 17-JAN-2002.  
PD  
XX  
XX 05-JUL-2001; 2001WO-EP07705.  
PF  
XX  
XX 07-JUL-2000; 2000DE-1033080.  
PR  
XX  
XX 19-APR-2001; 2001DE-1019294.  
PR  
XX  
XX (BOEH) BOEHRINGER INGELHEIM INT GMBH.  
PA  
XX  
XX  
PI  
XX  
XX

Schweifer N, Scherl-Mostageer M, Sommergruber W, Abseher R;  
DR MPI: 2002-171704/22.  
DR N-PSDB; ABA95507.  
XX  
XX New tumor-associated antigen B345, useful for diagnosis and  
PT immunotherapy of tumors, also related nucleic acid and antibodies -  
PT  
XX  
XX  
PS Claim 1; Page 85-88; 102pp; German.

XX This invention describes a novel tumour-associated antigen, designated  
CC B345 which has cytostatic activity. B345 is involved in communication,  
CC interaction and/or signal transduction with extracellular components and  
CC ligands, especially in the metastatic potential of cancers, particularly  
CC of the colon. B345 or its immunogenic fragments, also the DNA that  
CC encodes it, are useful for immunotherapy of cancer, particularly  
CC carcinoma of lung or colon. Antibodies raised against B345 are useful for  
CC treatment and diagnosis of cancers that are associated with B345  
CC expression, including their use for targeted delivery of cytotoxic or  
CC radioactive agents. Probes derived from B345 can be used to detect  
CC tumour-specific mutations in the B345 sequence, and can be used to screen  
CC for B345 specific modulators. This sequence represents a human B345  
CC tumour-associated antigen described in the invention.  
CC  
XX

Sequence 836 AA;



Query Match 100.0%; Score 836; DB 23; Length 836;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 836; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MAGLNCGVSIALLGVLLGGAARLPRGAFAPALPRESNTITLKLGTPTLAKCYIV 60
DB 1 MAGLNCGVSIALLGVLLGGAARLPRGAFAPALPRESNTITLKLGTPTLAKCYIV 60
QY 61 SKRHTMLSIKSERIVFTFSCQSPENHFVIEIOKNIDCMSPCPGEVQLOPSTSLPT 120
DB 61 SKRHTMLSIKSERIVFTFSCQSPENHFVIEIOKNIDCMSPCPGEVQLOPSTSLPT 120
QY 121 LNRFTIMDVKAHKSIGLELOFSIPRLROIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
DB 121 LNRFTIMDVKAHKSIGLELOFSIPRLROIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
QY 181 GTVSRIRKQOEGVKALHLPWFHPRNVSGFSIANRSSIRLCTIESVFEGEGSATLMSANY 240
DB 181 GTVSRIRKQOEGVKALHLPWFHPRNVSGFSIANRSSIRLCTIESVFEGEGSATLMSANY 240
QY 241 PEGPPEDELMTMOFVPAHLRASVSFLNPNLNCERKEEREVEYIPGSTNPEVFKLEDK 300
DB 241 PEGPPEDELMTMOFVPAHLRASVSFLNPNLNCERKEEREVEYIPGSTNPEVFKLEDK 300
QY 301 QPQNMAGNPNLSLQSCDDAOSPGILRLQFOVLVQHPONESNKIYVVDLSNERAMSLTIE 360
DB 301 QPQNMAGNPNLSLQSCDDAOSPGILRLQFOVLVQHPONESNKIYVVDLSNERAMSLTIE 360
QY 361 PRPVKQSKFVPGCVLCIESRTCSNLTLTSGSKKISFLCDLTRLMMNVEKTIISCTDH 420
DB 361 PRPVKQSKFVPGCVLCIESRTCSNLTLTSGSKKISFLCDLTRLMMNVEKTIISCTDH 420
QY 421 RYCQRKSYSLQVPSDILHLPEVLHDFSKMLVPRKORLSLVLPKQXLOQHTHEKCMNSF 480
DB 421 RYCQRKSYSLQVPSDILHLPEVLHDFSKMLVPRKORLSLVLPKQXLOQHTHEKCMNSF 480
QY 481 SYLVASAIPSODLYGSGFCPGGSIKQIOVKONISVTLRTFPAFQOEASROGLTVSFIPY 540
DB 481 SYLVASAIPSODLYGSGFCPGGSIKQIOVKONISVTLRTFPAFQOEASROGLTVSFIPY 540
QY 541 FKREGVFTVPTDYSKYVLTLPNMDRGLPSLTYSWNI SVPRDOVACLTFEKGSGVVCQ 600
DB 541 FKREGVFTVPTDYSKYVLTLPNMDRGLPSLTYSWNI SVPRDOVACLTFEKGSGVVCQ 600
QY 601 TGRAPMTIOEORTABEELFSLDEEDVLPRKSPFHHSFWNINISCSPTSGKQDLFLSVTLT 660
DB 601 TGRAPMTIOEORTABEELFSLDEEDVLPRKSPFHHSFWNINISCSPTSGKQDLFLSVTLT 660
QY 661 PRYVDLTVILIAAVGGVLLLSALGLIICVYKKKKKTKNGPAGVITNGNINTEMPROPK 720
DB 661 PRYVDLTVILIAAVGGVLLLSALGLIICVYKKKKKTKNGPAGVITNGNINTEMPROPK 720
QY 721 KFOKGRKQNDSHVAVIBEDTWYGHLLQDSSGSLFOPEVDTRAPQGTWGCPEPPTIC 780
DB 721 KFOKGRKQNDSHVAVIBEDTWYGHLLQDSSGSLFOPEVDTRAPQGTWGCPEPPTIC 780
QY 781 SRAPTAKLATEEPPRSPSESESEPTTSHPNNGVSSKODDIPILANTOEPEMPAE 836
DB 781 SRAPTAKLATEEPPRSPSESESEPTTSHPNNGVSSKODDIPILANTOEPEMPAE 836
```

RESULT 3  
AAM49640  
ID AAM49640 standard; Protein; 749 AA.

XX AAM49640;  
XX AC  
XX DT 17-MAY-2002 (first entry)  
DE Human tumour-associated antigen B345 protein.  
XX Tumour-associated antigen; human; B345; cytosolic; cell communication;

KW cell interaction; signal transduction; metastasis; cancer; colon;  
KW immunotherapy; carcinoma; lung; diagnosis.  
XX  
XX Homo sapiens.

```
QY 67 MLSTIKSERIVFTFSCQSPENHFVIEIOKNIDCMSPCPGEVQLOPSTSLPTLNRFTI 126
DB 1 MLSTIKSERIVFTFSCQSPENHFVIEIOKNIDCMSPCPGEVQLOPSTSLPTLNRFTI 126
QY 127 MDVKAHKSIGLELOFSIPRLROIGPESCPDGVTHSISGRIDATVVRIGTFCSNCTVSR 186
DB 127 MDVKAHKSIGLELOFSIPRLROIGPESCPDGVTHSISGRIDATVVRIGTFCSNCTVSR 186
QY 187 KMOEGVKAALHLPWFHPRNVSGFSIANRSSIRLCTIESVEGEGSATLMSANYPEGPPE 246
DB 187 KMOEGVKAALHLPWFHPRNVSGFSIANRSSIRLCTIESVEGEGSATLMSANYPEGPPE 246
QY 247 DELMTWQFVPAHLRASVSFLNPNLNCERKEEREVEYIPGSTNPEVFKLEDKQPGMA 306
DB 247 DELMTWQFVPAHLRASVSFLNPNLNCERKEEREVEYIPGSTNPEVFKLEDKQPGMA 306
QY 307 GNFNLSLQSCDDAOSPGILRLQFOVLVQHPONESNKIYVVDLSNERAMSLTIEPRPKQ 366
DB 307 GNFNLSLQSCDDAOSPGILRLQFOVLVQHPONESNKIYVVDLSNERAMSLTIEPRPKQ 366
QY 367 SRKFVPGCFVLCIESRTCSNLTLTSGSKKISFLCDLTRLMMNVEKTIISCTDHRVCOBK 426
DB 367 SRKFVPGCFVLCIESRTCSNLTLTSGSKKISFLCDLTRLMMNVEKTIISCTDHRVCOBK 426
QY 427 SYSLQVPSDILHLPEVLHDFSKMLVPRKORLSLVLPKQXLOQHTHEKCMNSFYLVAS 486
DB 427 SYSLQVPSDILHLPEVLHDFSKMLVPRKORLSLVLPKQXLOQHTHEKCMNSFYLVAS 486
```

QY 487 AIPSDLYGSGFCGGSIKQIQVKNISVTLRTFAPSFOQASRQGLTVSFIPEKEGV 546  
CC |||||  
CC AIPSDLYGSGFCGGSIKQIQVKNISVTLRTFAPSFOQASRQGLTVSFIPEKEGV 480  
DB 421 AIPSDLYGSGFCGGSIKQIQVKNISVTLRTFAPSFOQASRQGLTVSFIPEKEGV 480  
QY 547 FVTPTKRVKRVLRTPNMRGLPSLTLSVGMNITSVPDQVACLTFFERSGVCCQGRAM 606  
CC |||||  
CC FVTPTKRVKRVLRTPNMRGLPSLTLSVGMNITSVPDQVACLTFFERSGVCCQGRAM 540  
DB 481 FVTPTKRVKRVLRTPNMRGLPSLTLSVGMNITSVPDQVACLTFFERSGVCCQGRAM 540  
QY 607 IIOQRTRAEIIFSLDEVDLPKPSFHHHSFWMNISCSPTSGKQDLFSTVLTPTVDL 666  
CC |||||  
CC IIOQRTRAEIIFSLDEVDLPKPSFHHHSFWMNISCSPTSGKQDLFSTVLTPTVDL 600  
DB 541 IIOQRTRAEIIFSLDEVDLPKPSFHHHSFWMNISCSPTSGKQDLFSTVLTPTVDL 600  
QY 667 TVILAAVGGVLLSALGLIICVKKKKKTKNGPANGVINGNINTEMP 716  
CC |||||  
CC TVILAAVGGVLLSALGLIICVKKKKKTKNGPANGVINGNINTEMP 650  
DB 601 TVILAAVGGVLLSALGLIICVKKKKKTKNGPANGVINGNINTEMP 650

RESULT 4  
AAV91592  
ID AAV91592 standard; Protein; 443 AA.  
XX  
AC AAV91592;  
XX  
DT 29-JUN-2000 (first entry)  
XX  
DE Human secreted protein sequence encoded by gene 6 SEQ ID NO:265.  
XX  
KW Human; secreted protein; diagnosis; cytostatic; immunosuppressive;  
KW antiHIV; antiinflammatory; neurotropic; neuroprotective; antiallergic;  
KW osteopathic; antidiabetic; antibacterial; antidiabetic; antiasthma;  
KW antipneumonic; cardiant; gene therapy; cancer; neurological disorder;  
KW immune disease; inflammation; blood disorder; tumour.  
XX  
OS Homo sapiens.  
XX  
PN WO200006698-A1.  
XX  
PD 10-FEB-2000.  
XX  
PF 29-JUL-1999; 99WO-US17130.  
XX  
PR 30-JUL-1998; 98US-0094657.  
PR 05-AUG-1998; 98US-0095486.  
PR 06-AUG-1998; 98US-0095454.  
PR 06-AUG-1998; 98US-0095455.  
PR 12-AUG-1998; 98US-0096319.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Komatsoulis GA, Rosen CA, Ruben SM, Duan R, Moore PA, Shi Y;  
PI Lafleur D, Wei Y, Ni J, Florence KA, Young PE, Brewer LA;  
PI Soppet DR, Endress GA, Ebner R, Olsen HS, Mucenksi M;  
XX  
DR WPI; 2000-195282/17.  
XX  
PT New isolated human genes and the secreted polypeptides they encode,  
PT useful for diagnosis and treatment of e.g. cancers, neurological  
PT disorders, immune diseases, inflammation or blood disorders  
XX  
PS Disclosure; Page 18; 634pp; English.  
XX  
CC The polynucleotide sequences given in AAA26346 to AAA26458 encode the  
CC human secreted proteins given in AAV91451 to AAV91691. The human secreted  
CC proteins can have activities based on the tissues and cells they are  
CC expressed in. Examples of the activities are: cytostatic;  
CC immunosuppressive; antiHIV; antiinflammatory; neurotropic; neuroprotective;  
CC antiallergic; osteopathic; antidiabetic; antibacterial; antidiabetic;  
CC antiasthma; antipneumonic; cardiant and cancer. The polynucleotides and their  
CC corresponding secreted proteins are useful for preventing, treating or  
CC ameliorating medical conditions, e.g. by protein or gene therapy. Also  
CC pathological conditions can be diagnosed by determining the amount of the  
CC proteins in a sample or by determining the presence of mutations in the

CC polynucleotides. Specific uses are described for each of the  
CC polynucleotides, based on which tissues they are most highly expressed  
CC in, and include developing products for the diagnosis or treatment of  
CC cancer, tumours, neurodegenerative disorders, developmental abnormalities  
CC and foetal deficiencies, blood disorders, diseases of the immune system,  
CC autoimmune diseases, hepatic and renal disease, inflammation,  
CC osteoporosis, arthritis, infections, AIDS, spinal cord injuries,  
CC transplant rejection, diabetes, asthma, sepsis, acne, psoriasis,  
CC cardiovascular disorders, reproductive disorders, gastrointestinal  
CC disorders, respiratory disorders and metabolic disorders. The proteins  
CC or polynucleotides can also be used as food additives or preervatives.  
CC The proteins are also useful for identifying their binding partners.  
CC AAA26337 to AAA26345 and AAV91450 are sequences used in the  
CC exemplification of the present invention.  
XX  
SQ Sequence 443 AA;  
Query Match 49.4%; Score 413; DB 21; Length 443;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGLNCGVSIALLGLVLLGAARLPARGAFETALPRESNITVLIKGTPTLLAKPCYVI 60  
DB 30 MAGLNCGVSIALLGLVLLGAARLPARGAFETALPRESNITVLIKGTPTLLAKPCYVI 89  
QY 61 SKRHITMLSIKGERIVTFSCQSPENHFVIEIQKIDCMGSPCFGEVQLOPSTSLPT 120  
DB 90 SKRHITMLSIKGERIVTFSCQSPENHFVIEIQKIDCMGSPCFGEVQLOPSTSLPT 149  
QY 121 LNRFTMDVKAHKSIGLELOPSIPRLROIQGESCPDVTHSISGRIDATVYRIGTFCSN 180  
DB 150 LNRFTMDVKAHKSIGLELOPSIPRLROIQGESCPDVTHSISGRIDATVYRIGTFCSN 209  
QY 181 GTVSRIRKQEGVKALHLPWFPNPNVSGFSTANSTIRLCITIEVFEGSATLMSANY 240  
DB 210 GTVSRIRKQEGVKALHLPWFPNPNVSGFSTANSTIRLCITIEVFEGSATLMSANY 269  
QY 241 PEGFPEDELMTQFVVPALRLASVSLFNFNLSNCRKEKEVEYYPGSTNPEVEFKLBDK 300  
DB 270 PEGFPEDELMTQFVVPALRLASVSLFNFNLSNCRKEKEVEYYPGSTNPEVEFKLBDK 329  
QY 301 QPGNMAGFNLSLQCCDDAOSPGILRLQFOVLYOHPONESNKIYVDLSNERAMSLTE 360  
DB 330 QPGNMAGFNLSLQCCDDAOSPGILRLQFOVLYOHPONESNKIYVDLSNERAMSLTE 389  
QY 361 PRPVKQSRKFPVPGCFVCLSESTCSNLTLTSGSKHKISFLCDDTLRLMANYEK 413  
DB 390 PRPVKQSRKFPVPGCFVCLSESTCSNLTLTSGSKHKISFLCDDTLRLMANYEK 442

RESULT 5  
AAB95558  
ID AAB95558 standard; Protein; 649 AA.  
XX  
AC AAB95558;  
XX  
DT 26-JUN-2001 (first entry)  
XX  
DE Human protein sequence SEQ ID NO:18191.  
XX  
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.  
XX  
OS Homo sapiens.  
XX  
PN EP1074617-A2.  
XX  
PD 07-FEB-2001.  
XX  
PF 28-JUL-2000; 2000EP-0116126.  
XX  
PR 29-JUL-1999; 99JP-0248036.  
PR 27-AUG-1999; 99JP-0300253.

```

PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELI-) HELIX RES INST.
XX Ota T, Isegai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
PS Claim 8; SEQ ID 18191; 2537bp + CD ROM; English.
XX
CC The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 649 AA;
Query Match 45.7%; Score 382; DB 22; Length 649;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 582; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 253 QFVVPAAHLRASVSFLNPNLSNCRKEKEVEYIIPGTTNPEVFKLEDKQPGNMAGNFNLS 312
DB 66 QFVVPAAHLRASVSFLNPNLSNCRKEKEVEYIIPGTTNPEVFKLEDKQPGNMAGNFNLS 125
OY 313 LOGCDODQOSPGILARLOFVLOVHONSNKTYVVDLSNERAMSLTIPRPYKQSRKVP 372
DB 126 LOGCDODQOSPGILARLOFVLOVHONSNKTYVVDLSNERAMSLTIPRPYKQSRKVP 185
OY 373 GCFVCLSERTCSNLTLTSGSKHKISFLCDDTLRLMMVVEKTICTDHRVCORKSYSLOV 432
DB 186 GCFVCLSERTCSNLTLTSGSKHKISFLCDDTLRLMMVVEKTICTDHRVCORKSYSLOV 245
OY 433 PSDILHLPVELHDFSKLVLVPRDLRLVLVPAQKLQHTHEKPCNTSPSYLVASAPSD 492
DB 246 PGDILHLPVELHDFSKLVLVPRDLRLVLVPAQKLQHTHEKPCNTSPSYLVASAPSD 305
OY 493 LYFGSFCGSGSIKQIQVONKISVTLRTAPSFQASASQGLTVSFTIPFKKEGVTVTPD 552
DB 306 LYFGSFCGSGSIKQIQVONKISVTLRTAPSFQASASQGLTVSFTIPFKKEGVTVTPD 365
OY 553 TKSRYLTATPMMDRGLPSLTYSVNNISVPRDOVACLTFEKKERSGVVQCTGAFMIIOBOR 612
DB 366 TKSRYLTATPMMDRGLPSLTYSVNNISVPRDOVACLTFEKKERSGVVQCTGAFMIIOBOR 425
OY 613 TRAEIFSLDSDVLPKPSFHHHSFWNINSNCSPTSGKQDLDFSVTLTPRTVDLTVIIIA 672
DB 613 TRAEIFSLDSDVLPKPSFHHHSFWNINSNCSPTSGKQDLDFSVTLTPRTVDLTVIIIA 672

```

```

DB 426 TRAEIFSLDSDVLPKPSFHHHSFWNINSNCSPTSGKQDLDFSVTLTPRTVDLTVIIIA 485
OY 673 AVGGGVLLLSALGLIICVKKKKKKTKGPAVGILYNGNINTEMPROPKPKQGRKNDNH 732
DB 486 AVGGGVLLLSALGLIICVKKKKKKTKGPAVGILYNGNINTEMPROPKPKQGRKNDNH 545
OY 733 VYAVIEDTMVYGHLLQDSGSLQPEVDVTRPFGQTMGVCPPSPPTICRAPTAKLATEE 792
DB 546 VYAVIEDTMVYGHLLQDSGSLQPEVDVTRPFGQTMGVCPPSPPTICRAPTAKLATEE 605
OY 793 PPRSPSESESEPTFSHPNNGDVSSKOTDIPLNTQEPMEPAE 836
DB 606 PPRSPSESESEPTFSHPNNGDVSSKOTDIPLNTQEPMEPAE 649
RESULT 6
AAU12252
ID AAU12252 standard; Protein; 343 AA.
XX
AC AAU12252;
XX
DT 24-OCT-2001 (first entry)
XX
DE Human PRO5773 polypeptide sequence.
XX
KW Human secretory and transmembrane; PRO; mammalian; cancer; lung;
KW breast; prostate; cervical; tumor necrosis factor- $\alpha$ ; TNF- $\alpha$ ;
KW cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;
KW adipocyte; A-peptide; factor VIIa; gene therapy.
XX
OS Homo sapiens.
XX
PN MO200140466-A2.
XX
PD 07-JUN-2001.
XX
PF 01-DEC-2000; 2000WO-US32678.
XX
PR 01-DEC-1999; 99WO-US28301.
PR 01-DEC-1999; 99WO-US28634.
PR 02-DEC-1999; 99WO-US28551.
PR 02-DEC-1999; 99WO-US28564.
PR 02-DEC-1999; 99WO-US28565.
PR 09-DEC-1999; 99US-0170262.
PR 16-DEC-1999; 99WO-US30095.
PR 20-DEC-1999; 99WO-US30911.
PR 20-DEC-1999; 99WO-US30999.
PR 30-DEC-1999; 99WO-US31243.
PR 06-JAN-2000; 2000WO-US00277.
PR 06-JAN-2000; 2000WO-US00376.
PR 11-FEB-2000; 2000WO-US03565.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US04914.
PR 24-FEB-2000; 2000WO-US05004.
PR 01-MAR-2000; 2000WO-US05601.
PR 20-MAR-2000; 2000WO-US07377.
PR 21-MAR-2000; 2000WO-US07532.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 10-NOV-2000; 2000WO-US30873.
PA (GETH ) GENENTECH INC.
XX
PI Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W;
PI Gerliisen ME, Goddard A, Godowski PJ, Garney AL, Sherwood S;
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
XX
DR WPI; 2001-408281/43.

```

DR N-Peptide; AAS21324.  
XX Isolated, secretory and transmembrane PRO polypeptide used to detect  
PT other PRO polypeptides, link bioactive molecules to cells expressing  
PT PRO polypeptides, and detect the presence of mammalian tumours e.g.  
PT lung, breast, prostate, cervical -  
XX Claim 12; Fig 162; 813pp; English.  
PS  
XX AAU12172-AAU12446 represent novel human secretory and transmembrane  
CC PRO polypeptides. The PRO polypeptides are useful to detect other  
CC PRO polypeptides, to link bioactive molecules to cells expressing  
CC PRO polypeptides, to modulate biological activities of cells expressing  
CC PRO polypeptides, and to detect the presence of mammalian lung, colon,  
CC breast, prostate, rectal, cervical or liver tumours by comparing PRO  
CC polypeptide expression in a cell sample to that in a control sample.  
CC Some of the 275 sequences are also useful to stimulate the release of  
CC tumour necrosis factor-alpha (TNF-alpha) from human blood, the  
CC proliferation or differentiation of chondrocytes, the proliferation or  
CC gene expression in pericyte cells, the release of proteoglycans from  
CC cartilage, the proliferation of inner ear utricular supporting cells or  
CC of T-lymphocytes, the release of a cytokine from peripheral blood  
CC monocytes (PBMCs), or the proliferation of endothelial cells. Some of  
CC the PRO polypeptides may modulate glucose or free fatty acid uptake by  
CC skeletal muscle cells or by adipocytes, or inhibit binding of A-peptide  
CC to factor VIIa. The PRO polypeptides can be used in assays to identify  
CC molecules involved in binding interactions. The polynucleotides encoding  
CC PRO polypeptides can be used to generate probes, antisense RNA/DNA,  
CC transgenic or knock out animals and can be used in gene therapy.  
XX  
SQ Sequence 343 AA:  
Query Match 40.8%; Score 341; DB 22; Length 343;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MAGLNCGVSIALLGVLLGAAALPRGAEAFETALPRESNITVILIKIGTPTLAKCYIYI 60  
DB 1 MAGLNCGVSIALLGVLLGAAALPRGAEAFETALPRESNITVILIKIGTPTLAKCYIYI 60  
QY 61 SRKHITMLSIKSGERIVFTFSCSPENHVIETIKNIDMSGPCPGEVQLPSTSLPT 120  
DB 61 SRKHITMLSIKSGERIVFTFSCSPENHVIETIKNIDMSGPCPGEVQLPSTSLPT 120  
QY 121 LNRFTIMDKAKHSIGLELOFSIPRLRQIPGSESCPDGYTHSISGRIDATVVRIGTFCSN 180  
DB 121 LNRFTIMDKAKHSIGLELOFSIPRLRQIPGSESCPDGYTHSISGRIDATVVRIGTFCSN 180  
QY 181 GTVSRIKQOEGYKMAHLPEFHPRNVSIGFSIANRSSIKRLCTIESEYFEGSGATLMSANY 240  
DB 181 GTVSRIKQOEGYKMAHLPEFHPRNVSIGFSIANRSSIKRLCTIESEYFEGSGATLMSANY 240  
QY 241 PGPFPEDELMTQFVPAHLRASVSFLNENLSCNCKEKERVYIYI PGSTTNEVFLEBK 300  
DB 241 PGPFPEDELMTQFVPAHLRASVSFLNENLSCNCKEKERVYIYI PGSTTNEVFLEBK 300  
QY 301 QPGNAGNFNLSLOGCDQDQASPGILRLQFVLYVHPQNES 341  
DB 301 QPGNAGNFNLSLOGCDQDQASPGILRLQFVLYVHPQNES 341  
RESULT 7  
ABU66650  
ID ABU66650 standard; Protein; 343 AA.  
XX  
AC ABU66650;  
XX  
XX 23-MAY-2003 (first entry)  
XX  
XX Human PRO polypeptide #81.  
XX  
XX Human; PRO polypeptide; secreted and transmembrane protein;  
KW tumour necrosis factor-alpha; TNF-alpha; blood; proliferation;

KW differentiation; chondrocyte; tumour; genetic disorder;  
KW cytostatic.  
XX  
XX Homo sapiens.  
OS  
XX US2003036180-A1.  
PN  
XX 20-FEB-2003.  
PD  
XX  
XX  
PF 09-MAY-2002; 2002US-0143114.  
XX  
XX 31-MAR-1997; 97WO-US05230.  
PR 12-JUN-1998; 98WO-US12456.  
PR 14-JUL-1998; 98WO-US14552.  
PR 28-AUG-1998; 98WO-US17888.  
PR 10-SEP-1998; 98WO-US18824.  
PR 14-SEP-1998; 98WO-US19093.  
PR 14-SEP-1998; 98WO-US19094.  
PR 14-SEP-1998; 98WO-US19177.  
PR 16-SEP-1998; 98WO-US19330.  
PR 17-SEP-1998; 98WO-US19437.  
PR 07-OCT-1998; 98WO-US21141.  
PR 29-OCT-1998; 98WO-US22991.  
PR 29-OCT-1998; 98WO-US22992.  
PR 20-NOV-1998; 98WO-US24855.  
PR 01-DEC-1998; 98WO-US25108.  
PR 05-JAN-1999; 98WO-US25106.  
PR 08-MAR-1999; 98WO-US05028.  
PR 10-MAR-1999; 98WO-US05190.  
PR 20-APR-1999; 98WO-US08615.  
PR 14-MAY-1999; 98WO-US10733.  
PR 02-JUN-1999; 98WO-US12252.  
PR 01-SEP-1999; 98WO-US20111.  
PR 08-SEP-1999; 98WO-US20594.  
PR 13-SEP-1999; 98WO-US20944.  
PR 15-SEP-1999; 98WO-US21090.  
PR 15-SEP-1999; 98WO-US21547.  
PR 05-OCT-1999; 98WO-US23089.  
PR 29-NOV-1999; 98WO-US28214.  
PR 30-NOV-1999; 98WO-US28313.  
PR 30-NOV-1999; 98WO-US28409.  
PR 01-DEC-1999; 98WO-US28301.  
PR 01-DEC-1999; 98WO-US28634.  
PR 02-DEC-1999; 98WO-US28551.  
PR 02-DEC-1999; 98WO-US28564.  
PR 02-DEC-1999; 98WO-US28565.  
PR 16-DEC-1999; 98WO-US30095.  
PR 20-DEC-1999; 98WO-US30911.  
PR 20-DEC-1999; 98WO-US30999.  
PR 22-DEC-1999; 98WO-US30720.  
PR 30-DEC-1999; 98WO-US31243.  
PR 30-DEC-1999; 98WO-US31274.  
PR 05-JAN-2000; 2000WO-US00219.  
PR 06-JAN-2000; 2000WO-US00277.  
PR 06-JAN-2000; 2000WO-US00376.  
PR 11-FEB-2000; 2000WO-US03565.  
PR 18-FEB-2000; 2000WO-US04341.  
PR 18-FEB-2000; 2000WO-US04342.  
PR 22-FEB-2000; 2000WO-US04414.  
PR 24-FEB-2000; 2000WO-US04914.  
PR 24-FEB-2000; 2000WO-US05004.  
PR 01-MAR-2000; 2000WO-US05601.  
PR 02-MAR-2000; 2000WO-US05746.  
PR 02-MAR-2000; 2000WO-US05841.  
PR 10-MAR-2000; 2000WO-US06319.  
PR 15-MAR-2000; 2000WO-US06884.  
PR 20-MAR-2000; 2000WO-US07377.  
PR 21-MAR-2000; 2000WO-US07532.  
PR 30-MAR-2000; 2000WO-US08439.  
PR 17-MAY-2000; 2000WO-US13705.  
PR 22-MAY-2000; 2000WO-US14042.  
PR 30-MAY-2000; 2000WO-US14941.  
PR 02-JUN-2000; 2000WO-US15264.

28-JUL-2000; 2000WO-US20710.  
PR 11-AUG-2000; 2000WO-US22031.  
PR 23-AUG-2000; 2000WO-US23522.  
PR 24-AUG-2000; 2000WO-US23328.  
PR 08-NOV-2000; 2000WO-US30952.  
PR 10-NOV-2000; 2000WO-US30873.  
PR 01-DEC-2000; 2000WO-US34678.  
PR 20-DEC-2000; 2000WO-US34956.  
PR 28-FEB-2001; 2001WO-US06520.  
PR 01-MAR-2001; 2001WO-US06666.  
PR 25-MAY-2001; 2001WO-US17092.  
PR 01-JUN-2001; 2001WO-US17800.  
PR 20-JUN-2001; 2001WO-US19692.  
PR 22-JUN-2001; 2001WO-US20116.  
PR 29-JUN-2001; 2001WO-US21066.  
PR 09-JUL-2001; 2001WO-US21735.  
PR 20-DEC-2000; 2000US-0747259.  
PR 28-FEB-2001; 2001US-0796498.  
PR 09-MAR-2001; 2001US-0802706.  
PR 14-MAR-2001; 2001US-0808689.  
PR 22-MAR-2001; 2001US-0816744.  
PR 05-APR-2001; 2001US-0828366.  
PR 10-MAY-2001; 2001US-0854208.  
PR 10-MAY-2001; 2001US-0854280.  
PR 18-MAY-2001; 2001US-0860216.  
PR 25-MAY-2001; 2001US-0866028.  
PR 01-JUN-2001; 2001US-0872035.  
PR 05-JUN-2001; 2001US-0874503.  
PR 14-JUN-2001; 2001US-0882636.  
PR 19-JUN-2001; 2001US-0886342.  
PR 21-JUN-2001; 2001US-0887879.  
PR 18-JUL-2001; 2001US-0908827.  
PR 06-AUG-2001; 2001US-0924419.  
PR 09-AUG-2001; 2001US-0927796.  
PR 16-AUG-2001; 2001US-0931836.  
PR 19-DEC-2001; 2001US-0028072.  
  
(GETH ) GENENTECH INC.  
  
XX Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W,  
XX Gerltien ME, Goddard A, Godowski PJ, Gurney AU, Sherwood S,  
XX Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;  
XX MPI: 2003-332040/31.  
XX N-PSDB; ACN03683.  
  
XX New secreted and transmembrane PRO nucleic acids, useful for gene  
XX therapy, in chromosome and gene mapping, as chromosome markers, in  
XX tissue typing, and in chromosome identification -  
  
XX Claim 12; Fig 162; 660p; English.  
  
XX The present invention relates to the isolation of novel human PRO  
XX polypeptides, and the polynucleotide sequences encoding them. The  
XX PRO polypeptides are secreted and transmembrane proteins. The PRO  
XX polypeptides are useful for detecting other PRO polypeptides, for  
XX linking bioactive molecules to cells expressing PRO polypeptides,  
XX for modulating biological activities of cells expressing PRO  
XX polypeptides, and for identifying agonists or antagonists.  
XX The PRO polypeptides are useful for stimulating the release of  
XX tumour necrosis factor (TNF)-alpha from human blood, for stimulating  
XX the proliferation or differentiation of chondrocytes, and detecting the  
XX presence of tumours. The polynucleotide sequences encoding PRO  
XX polypeptides are useful as hybridisation probes, in chromosome and  
XX gene mapping, in the generation of antisense RNA and DNA, in the  
XX preparation of PRO polypeptides, for generating transgenic animals or  
XX knockout animals, for the genetic analysis of individuals with genetic  
XX disorders, and in gene therapy. AB06570-AB06684 represent the human  
XX PRO polypeptides of the invention.  
XX Note: The sequence data for this patent was obtained in electronic  
XX format directly from the USPTO web site at  
XX [seqdata.uspto.gov/patseqidentry.html](http://seqdata.uspto.gov/patseqidentry.html).

XX Sequence 343 AA;  
SQ Query Match 40.8%; Score 341; DB 24; Length 343;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MAGINCVSIALLVLLGAARLRGAEAFETALPRESNTIVLTKLGTPTLLARCYIVI 60  
DB 1 MAGINCVSIALLVLLGAARLRGAEAFETALPRESNTIVLTKLGTPTLLARCYIVI 60  
  
QY 61 SKRHITMISTKSGRIYFTFSCGSPENHFVLEIQXNDICMGPCPFGEVQLQPTSLIPT 120  
DB 61 SKRHITMISTKSGRIYFTFSCGSPENHFVLEIQXNDICMGPCPFGEVQLQPTSLIPT 120  
  
QY 121 LNRFTIMVVKAKHSGIGLEQPSIRLRIQGESCPDGVHSIGRIDATVVRIGTFCPSN 180  
DB 121 LNRFTIMVVKAKHSGIGLEQPSIRLRIQGESCPDGVHSIGRIDATVVRIGTFCPSN 180  
  
QY 181 GTVSRIRKQEGVKMALHPFHPNNVSGFSIANKSSIRKLCIISVFEGBSATLMSANY 240  
DB 181 GTVSRIRKQEGVKMALHPFHPNNVSGFSIANKSSIRKLCIISVFEGBSATLMSANY 240  
  
QY 241 PEGFPEDELMTWQVVPALPASVSLNPNLSNCRKEERYEYIIPGSTINPEVFKLEDK 300  
DB 241 PEGFPEDELMTWQVVPALPASVSLNPNLSNCRKEERYEYIIPGSTINPEVFKLEDK 300  
  
QY 301 QPGNMAGNPNLSLQGCDDQDQSPGILRLQFOVLVQHPONES 341  
DB 301 QPGNMAGNPNLSLQGCDDQDQSPGILRLQFOVLVQHPONES 341  
  
RESULT 8  
AB066926  
ID AB066926 standard; Protein: 343 AA.  
XX  
XX AB066926;  
XX  
XX 27-MAY-2003 (first entry)  
XX  
XX  
XX Human secreted/transmembrane, PRO, protein SEQ ID 162.  
XX  
XX  
XX Human; secreted protein; transmembrane protein; PRO;  
XX inflammatory disease; organ failure; atherosclerosis; cardiac injury;  
XX infertility; birth defects; premature aging; AIDS; biosensor;  
XX acquired immunodeficiency syndrome; cancer; diabetic complication;  
XX bioreactor; tumour.  
XX  
XX Homo sapiens.  
XX  
XX US2003032155-A1.  
XX  
XX 13-FEB-2003.  
XX  
XX 03-MAY-2002; 2002US-0137865.  
XX  
XX 31-MAR-1997; 97WO-US05230.  
XX 12-JUN-1998; 98WO-US12456.  
XX 14-JUL-1998; 98WO-US14552.  
XX 28-AUG-1998; 98WO-US17888.  
XX 10-SEP-1998; 98WO-US18824.  
XX 14-SEP-1998; 98WO-US19093.  
XX 14-SEP-1998; 98WO-US19094.  
XX 14-SEP-1998; 98WO-US19177.  
XX 16-SEP-1998; 98WO-US19330.  
XX 17-SEP-1998; 98WO-US19437.  
XX 07-OCT-1998; 98WO-US21141.  
XX 29-OCT-1998; 98WO-US22991.  
XX 29-OCT-1998; 98WO-US22992.  
XX 20-NOV-1998; 98WO-US24855.  
XX 01-DEC-1998; 98WO-US25108.  
XX 05-JAN-1999; 99WO-US00106.  
XX 08-MAR-1999; 99WO-US05028.

PR 10-MAR-1999; 99MO-US05190.  
PR 20-APR-1999; 99MO-US08615.  
PR 14-MAY-1999; 99MO-US10733.  
PR 02-JUN-1999; 99MO-US12252.  
PR 01-SEP-1999; 99MO-US20111.  
PR 08-SEP-1999; 99MO-US20594.  
PR 13-SEP-1999; 99MO-US20944.  
PR 15-SEP-1999; 99MO-US21090.  
PR 05-OCT-1999; 99MO-US21547.  
PR 29-NOV-1999; 99MO-US23089.  
PR 30-NOV-1999; 99MO-US28213.  
PR 30-NOV-1999; 99MO-US28313.  
PR 01-DEC-1999; 99MO-US28409.  
PR 01-DEC-1999; 99MO-US28301.  
PR 02-DEC-1999; 99MO-US28634.  
PR 02-DEC-1999; 99MO-US28551.  
PR 02-DEC-1999; 99MO-US28564.  
PR 02-DEC-1999; 99MO-US28565.  
PR 16-DEC-1999; 99MO-US30095.  
PR 20-DEC-1999; 99MO-US30911.  
PR 22-DEC-1999; 99MO-US30999.  
PR 30-DEC-1999; 99MO-US30720.  
PR 30-DEC-1999; 99MO-US31243.  
PR 05-JAN-2000; 99MO-US31274.  
PR 06-JAN-2000; 2000MO-US0219.  
PR 06-JAN-2000; 2000MO-US0277.  
PR 11-FEB-2000; 2000MO-US0376.  
PR 18-FEB-2000; 2000MO-US04365.  
PR 18-FEB-2000; 2000MO-US04341.  
PR 22-FEB-2000; 2000MO-US04342.  
PR 24-FEB-2000; 2000MO-US04414.  
PR 24-FEB-2000; 2000MO-US04914.  
PR 01-MAR-2000; 2000MO-US05004.  
PR 02-MAR-2000; 2000MO-US05601.  
PR 02-MAR-2000; 2000MO-US05746.  
PR 02-MAR-2000; 2000MO-US05841.  
PR 10-MAR-2000; 2000MO-US06119.  
PR 15-MAR-2000; 2000MO-US06884.  
PR 20-MAR-2000; 2000MO-US07377.  
PR 21-MAR-2000; 2000MO-US07532.  
PR 30-MAR-2000; 2000MO-US08439.  
PR 17-MAY-2000; 2000MO-US13705.  
PR 22-MAY-2000; 2000MO-US14042.  
PR 30-MAY-2000; 2000MO-US14941.  
PR 02-JUN-2000; 2000MO-US15264.  
PR 28-JUL-2000; 2000MO-US20710.  
PR 11-AUG-2000; 2000MO-US22031.  
PR 23-AUG-2000; 2000MO-US23522.  
PR 24-AUG-2000; 2000MO-US23528.  
PR 08-NOV-2000; 2000MO-US30952.  
PR 10-NOV-2000; 2000MO-US30873.  
PR 01-DEC-2000; 2000MO-US32678.  
PR 20-DEC-2000; 2000MO-US34856.  
PR 28-FEB-2001; 2001MO-US06520.  
PR 01-MAR-2001; 2001MO-US06666.  
PR 25-MAY-2001; 2001MO-US17092.  
PR 01-JUN-2001; 2001MO-US17800.  
PR 20-JUN-2001; 2001MO-US19692.  
PR 29-JUN-2001; 2001MO-US20116.  
PR 09-JUL-2001; 2001MO-US21066.  
PR 20-DEC-2000; 2000MO-US21735.  
PR 28-FEB-2001; 2001MO-US21759.  
PR 09-MAR-2001; 2001MO-US20706.  
PR 14-MAR-2001; 2001MO-US08689.  
PR 22-MAR-2001; 2001MO-US16744.  
PR 05-APR-2001; 2001MO-US28366.  
PR 10-MAY-2001; 2001MO-US4208.  
PR 18-MAY-2001; 2001MO-US4280.  
PR 25-MAY-2001; 2001MO-US46028.  
PR 01-JUN-2001; 2001MO-US46034.  
PR 01-JUN-2001; 2001MO-US46035.

PR 05-JUN-2001; 2001US-0874503.  
PR 14-JUN-2001; 2001US-0882636.  
PR 19-JUN-2001; 2001US-0886342.  
PR 21-JUN-2001; 2001US-0887879.  
PR 18-JUL-2001; 2001US-0908827.  
PR 06-AUG-2001; 2001US-0924419.  
PR 09-AUG-2001; 2001US-0927796.  
PR 16-AUG-2001; 2001US-0931836.  
PR 19-DEC-2001; 2001US-0028072.  
XX  
XX (GENETECH INC.  
XX Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W,  
XX Gerritsen ME, Goddard A, Godowski PJ, Gunney AL, Sherwood S,  
XX Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z,  
XX MPI: 2003-331925/31.  
XX N-PSDB; ACO4104.  
XX  
XX New secreted and transmembrane nucleic acids and polypeptides,  
XX designated as PRO, useful for treating inflammation, organ failure,  
XX atherosclerosis, cardiac injury, infertility, birth defects, premature  
XX aging, AIDS, or cancer -  
XX  
XX Claim 12; Fig 162; 659pp; English.  
XX  
XX The invention relates to an isolated nucleic acid comprising, or which is  
XX at least 80% identical to, or the full-length coding sequence of, any of  
XX the 275 nucleotide sequences, encoding the corresponding PRO polypeptide  
XX (one of 275 secreted or transmembrane proteins). The nucleic acid  
XX further comprises the full-length coding sequence of the DNA deposited  
XX under American Type Culture Collection (ATCC) accession number in a list  
XX cells in the specification. Also included are vectors and host  
XX cells for producing PRO proteins, PRO fusion proteins, anti-PRO  
XX antibodies, PRO extracellular domains and mature sequences, methods  
XX of detecting PRO proteins, methods for stimulating the release of  
XX TNF-alpha (tumour necrosis factor alpha) from human blood.  
XX (and the proliferation of differentiation of chondrocyte cells, the  
XX proliferation of, or gene expression in pericyte cells, the release or  
XX proteoglycans from cartilage, proliferation of inner ear utricular  
XX supporting cells, the proliferation of T-lymphocyte cells, the release  
XX of a cytokine from peripheral blood mononuclear cells (PBMC), or the  
XX proliferation of endothelial cells), a method for modulating the uptake  
XX of glucose or free fatty acid (FFA) by skeletal muscle cells,  
XX a method for inhibiting the binding of A-peptide to factor VIIA,  
XX or the differentiation of adipocyte cells, a method for detecting the  
XX presence of a tumour in a mammal and an oligonucleotide probe derived  
XX from any of the nucleotide sequences cited above. The nucleic acids and  
XX polypeptides are useful for treating inflammatory diseases, organ  
XX failure, atherosclerosis, cardiac injury, infertility, birth defects,  
XX premature aging, AIDS (acquired immunodeficiency syndrome), cancer, or  
XX diabetic complications. The nucleic acids are useful as hybridisation  
XX probes, in chromosome and gene mapping, and in generating antisense RNA  
XX or DNA. The polypeptides are useful as pharmaceuticals, diagnostics,  
XX bioassays or bioreactors. Both are useful in tissue typing.  
XX The present sequence represents a PRO protein of the invention.  
XX  
SQ Sequence 343 AA:  
Query Match 40.8%; Score 341; DB 24; Length 343;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MAGLNGVSIALGLVLLGAARLPGARFELPRESNITVLYKGPPTLAKPCYIY 60  
Db 1 MAGLNGVSIALGLVLLGAARLPGARFELPRESNITVLYKGPPTLAKPCYIY 60  
QY 61 SKRHITMLSIKSGERIVTFPSCSPENFVIEIQKIDMSGRCPCFGEVOLQPSLAPT 120  
Db 61 SKRHITMLSIKSGERIVTFPSCSPENFVIEIQKIDMSGRCPCFGEVOLQPSLAPT 120  
QY 121 INRTFIWVKAHKSIGLEQPSIPRLROIGPESCPDGVTHSISRIDATVVRIGTFCGN 180  
Db 121 INRTFIWVKAHKSIGLEQPSIPRLROIGPESCPDGVTHSISRIDATVVRIGTFCGN 180

Db 121 LNRTIWDVKAHKSIGLEIQSIPRLRQIGPESCPDGVTHISGRIDATVTRIGTFCSN 180  
 Qy 181 GTVSRIKQOEGVKMLHLFWPFRPNVSGFSIANRSSIKRLCTIEVSFEFGSATLMSANY 240  
 Db 181 GTVSRITKQOEGVKMLHLFWPFRPNVSGFSIANRSSIKRLCTIEVSFEFGSATLMSANY 240  
 Qy 241 PEGPPEDEIMTQFVVPALHRSVSGFLNLSNCRKEERVEYIPGTTNDEVEKLEDK 300  
 Db 241 PEGPPEDEIMTQFVVPALHRSVSGFLNLSNCRKEERVEYIPGTTNDEVEKLEDK 300  
 Qy 301 QPGNMAGNFNLISLQCGDDOASPGILRLQFVLVQHPPONES 341  
 Db 301 QPGNMAGNFNLISLQCGDDOASPGILRLQFVLVQHPPONES 341  
 RESULT 9  
 ABUS9731  
 ID ABUS9731 standard; Protein; 343 AA.  
 XX  
 AC ABUS9731;  
 XX  
 DT 13-MAY-2003 (first entry)  
 XX  
 DE Novel secreted and transmembrane protein PRO5773.  
 XX  
 KW Human; PRO; hypertrophy of neonatal heart; angiogenesis; wound healing;  
 KW cardiac insufficiency disorder; cancer; tumour; immune response;  
 KW adrenal cortical capillary endothelial growth; c-fos induction;  
 KW vascular endothelial growth factor inhibition; VEGF inhibition;  
 KW endothelial cell growth inhibitor; T-lymphocytes stimulation;  
 KW retinal neurons cell survival; rod photoreceptor cell survival;  
 KW retinal disorder; retinitis pigmentosum; kidney disorder;  
 KW mammalian kidney mesangial cell proliferation; Berger disease;  
 KW dermatitis; herpeticiformis; Cronh's disease; chondrocyte proliferation;  
 KW chondrocyte redifferentiation; sports injury; arthritis.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2003017563-A1.  
 XX  
 PD 23-JAN-2003.  
 XX  
 PF 07-MAY-2002; 2002US-0140808.  
 XX  
 XX 31-MAR-1997; 97WO-US05230.  
 PR 12-JUN-1998; 98WO-US12456.  
 PR 14-JUL-1998; 98WO-US14552.  
 PR 28-AUG-1998; 98WO-US17888.  
 PR 10-SEP-1998; 98WO-US18824.  
 PR 14-SEP-1998; 98WO-US19093.  
 PR 14-SEP-1998; 98WO-US19094.  
 PR 16-SEP-1998; 98WO-US19177.  
 PR 16-SEP-1998; 98WO-US19330.  
 PR 17-SEP-1998; 98WO-US19437.  
 PR 07-OCT-1998; 98WO-US21141.  
 PR 29-OCT-1998; 98WO-US22991.  
 PR 29-OCT-1998; 98WO-US22992.  
 PR 20-NOV-1998; 98WO-US24855.  
 PR 01-DEC-1998; 98WO-US25108.  
 PR 05-JAN-1999; 99WO-US00106.  
 PR 08-MAR-1999; 99WO-US05028.  
 PR 10-MAR-1999; 99WO-US05190.  
 PR 20-APR-1999; 99WO-US06615.  
 PR 14-MAY-1999; 99WO-US10733.  
 PR 02-JUN-1999; 99WO-US12252.  
 PR 01-SEP-1999; 99WO-US20111.  
 PR 08-SEP-1999; 99WO-US20594.  
 PR 13-SEP-1999; 99WO-US20944.  
 PR 15-SEP-1999; 99WO-US21090.  
 PR 05-SEP-1999; 99WO-US21547.  
 PR 05-OCT-1999; 99WO-US23089.  
 PR 29-NOV-1999; 99WO-US28214.  
 PR 30-NOV-1999; 99WO-US28313.

PR 30-NOV-1999; 99WO-US28409.  
 PR 01-DEC-1999; 99WO-US28301.  
 PR 01-DEC-1999; 99WO-US28634.  
 PR 02-DEC-1999; 99WO-US28551.  
 PR 02-DEC-1999; 99WO-US28564.  
 PR 02-DEC-1999; 99WO-US28565.  
 PR 16-DEC-1999; 99WO-US30095.  
 PR 20-DEC-1999; 99WO-US30911.  
 PR 22-DEC-1999; 99WO-US30999.  
 PR 30-DEC-1999; 99WO-US30720.  
 PR 30-DEC-1999; 99WO-US31243.  
 PR 30-DEC-1999; 99WO-US31274.  
 PR 05-JAN-2000; 2000WO-US00219.  
 PR 06-JAN-2000; 2000WO-US00277.  
 PR 06-JAN-2000; 2000WO-US00376.  
 PR 11-FEB-2000; 2000WO-US03565.  
 PR 18-FEB-2000; 2000WO-US04341.  
 PR 18-FEB-2000; 2000WO-US04342.  
 PR 22-FEB-2000; 2000WO-US04914.  
 PR 24-FEB-2000; 2000WO-US04914.  
 PR 24-FEB-2000; 2000WO-US05004.  
 PR 01-MAR-2000; 2000WO-US05601.  
 PR 02-MAR-2000; 2000WO-US05746.  
 PR 02-MAR-2000; 2000WO-US05841.  
 PR 10-MAR-2000; 2000WO-US06319.  
 PR 15-MAR-2000; 2000WO-US06884.  
 PR 20-MAR-2000; 2000WO-US07377.  
 PR 21-MAR-2000; 2000WO-US07537.  
 PR 30-MAR-2000; 2000WO-US08439.  
 PR 17-MAY-2000; 2000WO-US13705.  
 PR 22-MAY-2000; 2000WO-US14042.  
 PR 30-MAY-2000; 2000WO-US14941.  
 PR 02-JUN-2000; 2000WO-US15264.  
 PR 28-JUN-2000; 2000WO-US20710.  
 PR 11-AUG-2000; 2000WO-US22031.  
 PR 23-AUG-2000; 2000WO-US23522.  
 PR 24-AUG-2000; 2000WO-US23328.  
 PR 08-NOV-2000; 2000WO-US30952.  
 PR 10-NOV-2000; 2000WO-US30873.  
 PR 01-DEC-2000; 2000WO-US32678.  
 PR 20-DEC-2000; 2000WO-US34956.  
 PR 28-FEB-2001; 2001WO-US06520.  
 PR 01-MAR-2001; 2001WO-US06666.  
 PR 25-MAY-2001; 2001WO-US17092.  
 PR 01-JUN-2001; 2001WO-US17800.  
 PR 20-JUN-2001; 2001WO-US19692.  
 PR 22-JUN-2001; 2001WO-US20116.  
 PR 29-JUN-2001; 2001WO-US21735.  
 PR 09-JUL-2001; 2001WO-US21735.  
 PR 20-DEC-2000; 2000US-0747259.  
 PR 28-FEB-2001; 2001US-0796498.  
 PR 09-MAR-2001; 2001US-0802706.  
 PR 14-MAR-2001; 2001US-0808689.  
 PR 22-MAR-2001; 2001US-0816744.  
 PR 05-APR-2001; 2001US-0828366.  
 PR 10-MAY-2001; 2001US-0854208.  
 PR 10-MAY-2001; 2001US-0854280.  
 PR 18-MAY-2001; 2001US-0860216.  
 PR 25-MAY-2001; 2001US-0866028.  
 PR 01-JUN-2001; 2001US-0866034.  
 PR 05-JUN-2001; 2001US-0872035.  
 PR 14-JUN-2001; 2001US-0874503.  
 PR 14-JUN-2001; 2001US-0882636.  
 PR 19-JUN-2001; 2001US-0886342.  
 PR 21-JUN-2001; 2001US-0887879.  
 PR 18-JUL-2001; 2001US-0908827.  
 PR 06-AUG-2001; 2001US-0924419.  
 PR 09-AUG-2001; 2001US-0927796.  
 PR 16-AUG-2001; 2001US-0931836.  
 PR 19-DEC-2001; 2001US-0028072.  
 XX  
 XX  
 PA  
 XX (GETH ) GENENTECH INC.

PI Baker KP, Beresini M, DeForge L, Desnoyers L, Flyvbjerg E, Gao W;  
PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;  
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;  
DR WPI: 2003-148238/14.  
DR N-PSDB; ABR99221.  
XX  
XX  
PS Claim 12; Fig 162; 659pp; English.  
XX  
XX The invention describes an isolated human PRO polypeptide. The PRO  
CC polypeptides are useful in detecting PRO polypeptides in a sample, in  
CC linking a bioactive molecule to a cell expressing a PRO polypeptide, and  
CC in modulating at least one biological activity of a cell expressing a PRO  
CC polypeptide. PRO1312 stimulates hypertrophy of neonatal heart and is thus  
CC useful for treating cardiac insufficiency disorders. PRO1154 and PRO1186  
CC stimulate adrenal cortical capillary endothelial growth, and PRO536,  
CC PRO943, PRO828, PRO1068 or PRO535, PRO826, PRO819, PRO1126,  
CC PRO1360 and PRO1387 induce c-fos in endothelial cells, and are thus  
CC useful for treating conditions or disorders where angiogenesis would be  
CC beneficial, e.g. wound healing and angiogenesis of this polypeptide are  
CC useful for treating cancerous tumours. PRO812 inhibits vascular  
CC endothelial growth factor (VEGF) stimulated proliferation of endothelial  
CC cells and is thus useful for inhibiting endothelial cell growth in  
CC mammals which would be beneficial in inhibiting tumour growth. PRO826,  
CC PRO1068, PRO1184, PRO1346 and PRO1375 stimulate proliferation of  
CC stimulated T-lymphocytes and are therapeutically useful for enhancing  
CC immune response. PRO828, PRO826, PRO1068 or PRO1132 enhance survival of  
CC retinal neurons cells (PRO1132 is also enhances survival/proliferation of  
CC retinal photoreceptor cells) and therefore are useful for treating retinal  
CC disorders of injuries, e.g. retinitis pigmentosa, AMD, PRO819, PRO813  
CC and PRO1066 induce proliferation of mammalian kidney mesangial cells,  
CC and therefore are useful for treating kidney disorders associated with  
CC decreased mesangial cell function such as Berger disease or other  
CC nephropathies associated with dermatitis, herpeticiformis or Crohn's  
CC disease. PRO1310, PRO844, PRO1312, PRO1192 and PRO1387 induce the  
CC proliferation and/or redifferentiation of chondrocytes in culture and  
CC are thus useful for treating sports injuries and arthritis. This  
CC is the amino acid sequence of a novel human PRO protein.  
XX  
XX Sequence 343 AA:  
SQ  
Query Match. 40.8%; Score 341; DB 24; Length 343;  
Best Local Similarity 100.0%; Pred No. 0;  
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MAGLNGVSIALGLVLLGAAALPRGAFAEIALPRESNTVLIKLTPTLAKPCYIV 60  
DB 1 MAGLNGVSIALGLVLLGAAALPRGAFAEIALPRESNTVLIKLTPTLAKPCYIV 60  
QY 61 SKRHITMLSTKSGERIVFTFSCSPENHPIETIQKIDCMGSGCPGEVQLPSTSLPT 120  
DB 61 SKRHITMLSTKSGERIVFTFSCSPENHPIETIQKIDCMGSGCPGEVQLPSTSLPT 120  
QY 121 LNRTFTMDVKAHSGISGLTLPSTLRLOIGPESGCDGVTHTSGIDATVVRIGFGCN 180  
DB 121 LNRTFTMDVKAHSGISGLTLPSTLRLOIGPESGCDGVTHTSGIDATVVRIGFGCN 180  
QY 181 GTVSRITKMGEGVMALHLPMFHPRNVSGPSIANRSSIKRLCTIESVFEBSGATLSANY 240  
DB 181 GTVSRITKMGEGVMALHLPMFHPRNVSGPSIANRSSIKRLCTIESVFEBSGATLSANY 240  
QY 241 PEGFPEDELMTQFVPAHLRAVSFLNFTLSCERKEVEREYITGTTNPEVFLDEK 300  
DB 241 PEGFPEDELMTQFVPAHLRAVSFLNFTLSCERKEVEREYITGTTNPEVFLDEK 300  
QY 301 QPNNMAGNFNLSLOGDDDAOSGIIIRLOPQVVOHPONES 341  
DB 301 QPNNMAGNFNLSLOGDDDAOSGIIIRLOPQVVOHPONES 341

RESULT 10  
AAV91456  
ID AAV91456 standard; Protein; 709 AA.  
XX  
XX AAV91456;  
AC  
XX 29-JUN-2000 (first entry)  
DT  
XX  
XX Human secreted protein sequence encoded by gene 6 SEQ ID NO:129.  
DE  
XX  
XX Human; secreted protein; diagnosis; cytostatic; immunosuppressive;  
KW antiHIV; antiinflammatory; nootropic; neuroprotective; antiallergic;  
KW osteopathic; antiautistic; antibacterial; antidiabetic; antiashma;  
KW antipsoriatic; cardiant; gene therapy; cancer; neurological disorder;  
KW immune disease; inflammation; blood disorder; tumour.  
XX  
OS Homo sapiens.  
XX  
XX WO200006698-A1.  
XX 10-FEB-2000.  
XX  
XX 29-JUL-1999; 99WO-US17130.  
XX  
XX 30-JUL-1998; 98US-0094657.  
XX 05-AUG-1998; 98US-0095486.  
XX 06-AUG-1998; 98US-0095454.  
XX 06-AUG-1998; 98US-0095455.  
XX 12-AUG-1998; 98US-0096319.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Komatsuji GA, Rosen CA, Ruben SM, Duan R, Moore PA, Shi Y;  
PI Lafleur D, Wei Y, Ni J, Florense KA, Young PE, Brewer LA;  
PI Soppet DR, Endres GA, Emner R, Olsen HS, Mucenski M;  
XX WPI: 2000-195282/17.  
XX N-PSDB; AAA26351.  
XX  
XX New isolated human genes and the secreted polypeptides they encode,  
PT useful for diagnosis and treatment of e.g. cancer, neurological  
PT disorders, immune diseases, inflammation or blood disorders -  
XX  
XX Claim 11; Page 456-459; 634pp; English.  
PS  
XX The polynucleotide sequences given in AAA26346 to AAA26458 encode the  
CC human secreted proteins given in AAV91451 to AAV91461. The human secreted  
CC proteins can have activities based on the tissues and cells they are  
CC expressed in. Examples of the activities are: cytostatic;  
CC immunosuppressive; antiHIV; antiinflammatory; nootropic; neuroprotective;  
CC antiallergic; osteopathic; antiautistic; antibacterial; antidiabetic;  
CC antiashma; antipsoriatic; and cardiant. The polynucleotides and their  
CC corresponding secreted proteins are useful for preventing, treating or  
CC ameliorating medical conditions, e.g. by protein or gene therapy. Also  
CC pathological conditions can be diagnosed by determining the amount of the  
CC proteins in a sample or by determining the presence of mutations in the  
CC polynucleotides. Specific uses are described for each of the  
CC polynucleotides, based on which tissues they are most highly expressed  
CC in, and include developing products for the diagnosis or treatment of  
CC cancer, tumours, neurodegenerative disorders, developmental abnormalities  
CC and foetal deficiencies, blood disorders, diseases of the immune system,  
CC autoimmune diseases, hepatic and renal diseases, inflammation,  
CC allergies, Alzheimer's and behavioural disorders, schizophrenia,  
CC osteoporosis, arthritis, infections, AIDS, spinal cord injuries,  
CC transplant rejection, diabetes, asthma, sepsis, acne, psoriasis,  
CC cardiovascular disorders, reproductive disorders, gastrointestinal  
CC disorders, respiratory disorders and metabolic disorders. The proteins  
CC or polynucleotides can also be used as food additives or preervatives.  
CC The proteins are also useful for identifying their binding partners.  
CC AAA26337 to AAA26345 and AAV91450 are sequences used in the  
CC exemplification of the present invention.



```

XX SQ Sequence 709 AA:
Query Match 25.2%; Score 211; DB 21; Length 709;
Best Local Similarity 99.5%; Pred. No. 2,9e-197;
Matches 411; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 MAGLNCGVSIALLGVLLGAARLPGAAEAFIALPRESNITVLIKLGPPTLAKPCYIV 60
DB 1 MAGLNCGVSIALLGVLLGAARLPGAAEAFIALPRESNITVLIKLGPPTLAKPCYIV 60
OY 61 SKRHITMLSIKSGERIVFTFSCQSPBNHFVITIQKNIDCMSPCPGGEVQLPSTSLPT 120
DB 61 SKRHITMLSIKSGERIVFTFSCQSPBNHFVITIQKNIDCMSPCPGGEVQLPSTSLPT 120
OY 121 LNRITIMDKAKHSIGLELOFSIPRLROIQGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
DB 121 LNRITIMDKAKHSIGLELOFSIPRLROIQGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
OY 181 GTVSRIKMGEGVMALHLPMPHPRNVSGFSIANRSSIKLCITIEVFEGEGSATLMSANY 240
DB 181 GTVSRIKMGEGVMALHLPMPHPRNVSGFSIANRSSIKLCITIEVFEGEGSATLMSANY 240
OY 241 PEGFPEDELMTWQFVVPAPHLRASVSFLNPNLSNCERKEERVEYIIPGSTTNDPFLBDK 300
DB 241 PEGFPEDELMTWQFVVPAPHLRASVSFLNPNLSNCERKEERVEYIIPGSTTNDPFLBDK 300
OY 301 OPGNAGNPNLSLQCCDDOAGSPGILRLQFOVLVQHPOESNKIYVVDLSNRRAMSLTTE 360
DB 301 OPGNAGNPNLSLQCCDDOAGSPGILRLQFOVLVQHPOESNKIYVVDLSNRRAMSLTTE 360
OY 361 PRPVQSRKRVPGCFVCLSESRFCSSNLTLSGSKHISFLCDLTRLMMNVK 413
DB 361 PRPVQSRKRVPGCFVCLSESRFCSSNLTLSGSKHISFLCDLTRLMMNVK 413

RESULT 11
ID AAY91593 standard; Protein; 159 AA.
AC AAY91593;
XX
XX 29-JUN-2000 (first entry)
XX
DE Human secreted protein sequence encoded by gene 6 SEQ ID NO:266.
XX
XX Human; secreted protein; diagnosis; cytostatic; immunosuppressive;
XX anti-HIV; antiinflammatory; nootropic; neuroprotective; antiasthma;
XX osteopathic; antiarthritic; antibacterial; antidiabetic; antiashma;
XX antiparasitic; cardiant; gene therapy; cancer; neurological disorder;
XX immune disease; inflammation; blood disorder; tumour.
XX
XX Homo sapiens.
XX
XX WO200006698-A1.
XX
XX 10-FEB-2000.
XX
XX 29-JUL-1999; 99WO-US17130.
XX
XX 30-JUL-1998; 98US-0094657.
XX 05-AUG-1998; 98US-0095486.
XX 06-AUG-1998; 98US-0095454.
XX 06-AUG-1998; 98US-0095455.
XX 12-AUG-1998; 98US-0096319.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Komatsugilis GA, Rosen CA, Ruben SM, Duan R, Moore PA, Shi Y;
XX Lafleur D, Wei Y, Ni J, Florence KA, Young PE, Brewer LA;
XX Soppet DR, Endreess GA, Ebner R, Olsen HS, Mucenski M;
XX
XX WPI; 2000-195282/17.

```

```

XX New isolated human genes and the secreted polypeptides they encode,
XX useful for diagnosis and treatment of e.g. cancers, neurological
XX disorders, immune diseases, inflammation or blood disorders -
XX
XX Disclosure; Page 18; 634pp; English.
XX
XX The polynucleotide sequences given in AAA26346 to AAA26458 encode the
XX human secreted proteins given in AAY91451 to AAY91691. The human secreted
XX proteins can have activities based on the tissues and cells they are
XX expressed in. Examples of the activities are: cytostatic;
XX immunosuppressive; anti-HIV; antiinflammatory; nootropic; neuroprotective;
XX antiasthma; osteopathic; antiarthritic; antibacterial; antidiabetic;
XX antiparasitic; cardiant; and cardiant. The polynucleotides and their
XX corresponding secreted proteins are useful for preventing, treating or
XX ameliorating medical conditions, e.g. by protein or gene therapy. Also
XX pathological conditions can be diagnosed by determining the amount of the
XX proteins in a sample or by determining the presence of mutations in the
XX polynucleotides. Specific uses are described for each of the
XX polynucleotides, based on which tissues they are most highly expressed
XX in, and include developing products for the diagnosis or treatment of
XX cancer, tumours, neurodegenerative disorders, developmental abnormalities
XX and foetal deficiencies, blood disorders, diseases of the immune system,
XX autoimmune diseases, hepatic and renal disease, inflammation,
XX allergies, Alzheimer's and behavioural disorders, schizophrenia,
XX osteoporosis, arthritis, infections, AIDS, spinal cord injuries,
XX transplant rejection, diabetes, asthma, sepsis, acne, psoriasis,
XX cardiovascular disorders, reproductive disorders, gastrointestinal
XX disorders, respiratory disorders and metabolic disorders. The proteins
XX or polynucleotides can also be used as food additives or preservatives.
XX The proteins are also useful for identifying their binding partners.
XX AAA26346 to AAA26345 and AAY91450 are sequences used in the
XX exemplification of the present invention.
XX
XX Sequence 159 AA:
Query Match 19.0%; Score 159; DB 21; Length 159;
Best Local Similarity 100.0%; Pred. No. 6,3e-147;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 30 FEIALPRESNITVLIKLGPPTLAKPCYIVISKRHITMLSIKSGRIIVFTFSCQSPBNHF 89
DB 1 FEIALPRESNITVLIKLGPPTLAKPCYIVISKRHITMLSIKSGRIIVFTFSCQSPBNHF 89
OY 90 VIEIQKNIDCMSPCPGGEVQLPSTSLPTLNRFIMDVAKHSIGLELOFSIPRLROI 149
DB 61 VIEIQKNIDCMSPCPGGEVQLPSTSLPTLNRFIMDVAKHSIGLELOFSIPRLROI 120
OY 150 PGESCPDGVTHSISGRIDATVVRIGTFCSNGTVSRIM 188
DB 121 PGESCPDGVTHSISGRIDATVVRIGTFCSNGTVSRIM 159

RESULT 12
ID ABG00296 standard; Protein; 143 AA.
AC ABG00296;
XX
XX 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #287.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX Homo sapiens.
XX
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US08631.

```

XX 31-MAR-2000; 2000US-0540217.  
PR 23-AUG-2000; 2000US-0649167.  
XX (HYSE-) HYSEQ INC.  
XX Drmanac RT, Liu C, Tang YT;  
PI WPI, 2001-639362/73.  
XX N-PSDB; AAS64483.  
XX  
XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensic, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity -  
XX  
XX Claim 20; SEQ ID No 30655; 103pp; English.  
XX  
XX The invention relates to isolated polynucleotide (I) and  
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensic, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABG30377 represent novel human  
CC diagnostic amino acid sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX Sequence 143 AA;  
SQ  
Query Match 17.1%; Score 143; DB 22; Length 143;  
Best Local Similarity 100.0%; Pred. No. 2.6e-131;  
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 694 KKKKTKGPAVGIYNGNINTEMPOPKKQKGRKNDSHVYAVIEDTMYGHLQDSGS 753  
DB 1 KKKKTKGPAVGIYNGNINTEMPOPKKQKGRKNDSHVYAVIEDTMYGHLQDSGS 60  
QY 754 FLOPEVDYTRPFQGTGVCPPSPPTICSRAPAKLATEEPPRSPSESESEPTFSHPNN 813  
DB 61 FLOPEVDYTRPFQGTGVCPPSPPTICSRAPAKLATEEPPRSPSESESEPTFSHPNN 120  
QY 814 GDVSSKDTDIPILNTOEPMEPAE 836  
DB 121 GDVSSKDTDIPILNTOEPMEPAE 143  
RESULT 13  
AAM23969  
ID AAM23969 standard; Protein; 121 AA.  
XX  
XX AAM23969;  
AC  
XX  
XX 12-OCT-2001 (first entry)  
DT  
XX  
XX Human EST encoded protein SEQ ID NO: 1494.  
DE  
XX  
XX Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;  
KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;  
KW diagnostics; forensic test; gene mapping; genetic disorder;  
KW biodiversity; gene therapy; nutrition.

XX Homo sapiens.  
OS  
XX  
XX WO200154477-A2.  
PN  
XX  
XX 02-AUG-2001.  
PD  
XX  
XX 25-JAN-2001; 2001WO-US02687.  
PF  
XX  
XX 25-JAN-2000; 2000US-0491404.  
PR 17-JUL-2000; 2000US-0617746.  
PR 03-AUG-2000; 2000US-0631451.  
PR 15-SEP-2000; 2000US-0663870.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
XX Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;  
PI Cao Y, Drmanac RA, Zhang J, Werhman T;  
XX N-PSDB; AAH98628.  
DR  
XX WPI, 2001-476164/51.  
XX  
XX Isolated polypeptide for treatment of diseases, diagnostics, raising  
PT antibodies and research use -  
XX  
XX Claim 20; Page 1036; 1275pp; English.  
XX  
XX The present invention provides the protein and coding sequences of novel  
CC proteins from a variety of organisms, including human, dog, cat, horse,  
CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea  
CC urchin and tomato. These were derived from expressed sequence tags (ESTs)  
CC from the organism of interest. They can be used in diagnostics,  
CC forsnasics, gene mapping, identification of mutations, to assess  
CC biodiversity and for nutritional purposes. The present sequence is a  
CC protein of the invention.  
XX  
XX Sequence 121 AA;  
SQ  
Query Match 14.5%; Score 121; DB 22; Length 121;  
Best Local Similarity 100.0%; Pred. No. 7.9e-110;  
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MAGLNCGVSIALLGVLGAAALPRGAFERIALPRESNITVLYLGTPTLAKCYIYI 60  
DB 1 MAGLNCGVSIALLGVLGAAALPRGAFERIALPRESNITVLYLGTPTLAKCYIYI 60  
QY 61 SKRHITMLSIKSGRIVTFSCQSPENHFVIEIOKNIDCMGSGPCFGEVOLQPSLLPT 120  
DB 61 SKRHITMLSIKSGRIVTFSCQSPENHFVIEIOKNIDCMGSGPCFGEVOLQPSLLPT 120  
QY 121 L 121  
DB 121 L 121  
RESULT 14  
AAM90342  
ID AAM90342 standard; Protein; 116 AA.  
XX  
XX AAM90342;  
AC  
XX  
XX 07-NOV-2001 (first entry)  
DT  
XX  
XX Human immune/haematopoietic antigen SEQ ID NO:17935.  
DE  
XX  
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
KW cytostatic; gene therapy; vaccine; metastasis.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO200157182-A2.  
PN  
XX  
XX 09-AUG-2001.

XX 17-JAN-2001; 2001WO-US01354.  
 PF 31-JAN-2000; 2000US-0179065.  
 XX 04-FEB-2000; 2000US-0180628.  
 PR 04-FEB-2000; 2000US-0184664.  
 PR 02-MAR-2000; 2000US-0186350.  
 PR 16-MAR-2000; 2000US-0189874.  
 PR 17-MAR-2000; 2000US-0190076.  
 PR 18-APR-2000; 2000US-0198123.  
 PR 19-MAY-2000; 2000US-0205515.  
 PR 07-JUN-2000; 2000US-0209467.  
 PR 28-JUN-2000; 2000US-0214886.  
 PR 30-JUN-2000; 2000US-0215135.  
 PR 07-JUL-2000; 2000US-0216647.  
 PR 07-JUL-2000; 2000US-0216880.  
 PR 11-JUL-2000; 2000US-0217487.  
 PR 11-JUL-2000; 2000US-0217496.  
 PR 14-JUL-2000; 2000US-0218290.  
 PR 26-JUL-2000; 2000US-0220963.  
 PR 26-JUL-2000; 2000US-0220964.  
 PR 14-AUG-2000; 2000US-0224518.  
 PR 14-AUG-2000; 2000US-0224519.  
 PR 14-AUG-2000; 2000US-0225213.  
 PR 14-AUG-2000; 2000US-0225214.  
 PR 14-AUG-2000; 2000US-0225266.  
 PR 14-AUG-2000; 2000US-0225267.  
 PR 14-AUG-2000; 2000US-0225268.  
 PR 14-AUG-2000; 2000US-0225270.  
 PR 14-AUG-2000; 2000US-0225447.  
 PR 14-AUG-2000; 2000US-0225757.  
 PR 14-AUG-2000; 2000US-0225758.  
 PR 14-AUG-2000; 2000US-0225759.  
 PR 18-AUG-2000; 2000US-0225729.  
 PR 22-AUG-2000; 2000US-0226811.  
 PR 22-AUG-2000; 2000US-0226868.  
 PR 23-AUG-2000; 2000US-0227182.  
 PR 30-AUG-2000; 2000US-0228924.  
 PR 01-SEP-2000; 2000US-0228287.  
 PR 01-SEP-2000; 2000US-0229343.  
 PR 01-SEP-2000; 2000US-0229344.  
 PR 01-SEP-2000; 2000US-0229345.  
 PR 05-SEP-2000; 2000US-0229509.  
 PR 05-SEP-2000; 2000US-0229513.  
 PR 06-SEP-2000; 2000US-0230437.  
 PR 06-SEP-2000; 2000US-0230438.  
 PR 08-SEP-2000; 2000US-0231242.  
 PR 08-SEP-2000; 2000US-0231243.  
 PR 08-SEP-2000; 2000US-0231244.  
 PR 08-SEP-2000; 2000US-0231413.  
 PR 08-SEP-2000; 2000US-0231414.  
 PR 08-SEP-2000; 2000US-0232080.  
 PR 08-SEP-2000; 2000US-0232081.  
 PR 12-SEP-2000; 2000US-0231968.  
 PR 14-SEP-2000; 2000US-0233397.  
 PR 14-SEP-2000; 2000US-0233398.  
 PR 14-SEP-2000; 2000US-0233399.  
 PR 14-SEP-2000; 2000US-0233400.  
 PR 14-SEP-2000; 2000US-0233401.  
 PR 14-SEP-2000; 2000US-0233063.  
 PR 14-SEP-2000; 2000US-0233064.  
 PR 14-SEP-2000; 2000US-0233065.  
 PR 21-SEP-2000; 2000US-0234223.  
 PR 21-SEP-2000; 2000US-0234274.  
 PR 25-SEP-2000; 2000US-0234997.  
 PR 25-SEP-2000; 2000US-0234998.  
 PR 26-SEP-2000; 2000US-0235484.  
 PR 27-SEP-2000; 2000US-0235834.  
 PR 27-SEP-2000; 2000US-0235836.  
 PR 29-SEP-2000; 2000US-0236327.  
 PR 29-SEP-2000; 2000US-0236367.  
 PR 29-SEP-2000; 2000US-0236368.

PR 29-SEP-2000; 2000US-0236369.  
 PR 29-SEP-2000; 2000US-0236370.  
 PR 02-OCT-2000; 2000US-0236802.  
 PR 02-OCT-2000; 2000US-0237037.  
 PR 02-OCT-2000; 2000US-0237038.  
 PR 02-OCT-2000; 2000US-0237039.  
 PR 02-OCT-2000; 2000US-0237040.  
 PR 13-OCT-2000; 2000US-0239935.  
 PR 13-OCT-2000; 2000US-0239937.  
 PR 20-OCT-2000; 2000US-0240960.  
 PR 20-OCT-2000; 2000US-0241221.  
 PR 20-OCT-2000; 2000US-0241785.  
 PR 20-OCT-2000; 2000US-0241786.  
 PR 20-OCT-2000; 2000US-0241787.  
 PR 20-OCT-2000; 2000US-0241808.  
 PR 20-OCT-2000; 2000US-0241809.  
 PR 01-NOV-2000; 2000US-0241826.  
 PR 08-NOV-2000; 2000US-0246474.  
 PR 08-NOV-2000; 2000US-0246475.  
 PR 08-NOV-2000; 2000US-0246476.  
 PR 08-NOV-2000; 2000US-0246477.  
 PR 08-NOV-2000; 2000US-0246478.  
 PR 08-NOV-2000; 2000US-0246523.  
 PR 08-NOV-2000; 2000US-0246524.  
 PR 08-NOV-2000; 2000US-0246525.  
 PR 08-NOV-2000; 2000US-0246526.  
 PR 08-NOV-2000; 2000US-0246527.  
 PR 08-NOV-2000; 2000US-0246528.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246609.  
 PR 08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 08-NOV-2000; 2000US-0246613.  
 PR 17-NOV-2000; 2000US-0249207.  
 PR 17-NOV-2000; 2000US-0249208.  
 PR 17-NOV-2000; 2000US-0249209.  
 PR 17-NOV-2000; 2000US-0249210.  
 PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249244.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249264.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249299.  
 PR 01-DEC-2000; 2000US-0250300.  
 PR 01-DEC-2000; 2000US-0250310.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 05-DEC-2000; 2000US-0256719.  
 PR 06-DEC-2000; 2000US-0256719.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251988.  
 PR 08-DEC-2000; 2000US-0251990.  
 PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0259678.  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX Rosen CA, Barash SC, Ruben SM,  
 XX WPI, 2001-483426/52.  
 DR N-PSDB; AAK63123.  
 DR

XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
PT useful for preventing, diagnosing and/or treating cancers and  
PT metastasis -  
XX  
XX Claim 11: SEQ ID NO 17935; 3071bp + Sequence Listing; English.  
XX  
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)  
CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic  
CC activity, and can be given in gene therapy and vaccine production. (I)  
CC proteins and polynucleotides may be used in the prevention, diagnosis and  
CC treatment of diseases associated with inappropriate (I) expression. For  
CC example, they may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of (I) by expressing inactive proteins or to  
CC supplement the patient's own production of (I). Additionally, (I)  
CC polynucleotides may be used to produce the secreted (I), by inserting  
CC the nucleic acids into a host cell and culturing the cell to express the  
CC protein. (I) proteins and polynucleotides may be used to prevent,  
CC diagnose and treat immune/hematopoietic-related diseases, especially  
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703  
CC to AAK87694 represent human immune/hematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169  
CC represent sequences used in the exemplification of the present invention.  
XX  
XX Sequence 116 AA;  
SQ  
Query Match 8.9%; Score 74; DB 22; Length 116;  
Best Local Similarity 100.0%; Pred. No. 7.8e-64;  
Matches 74; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 343 KIYVVDLSNERAMSLTIEPRPVKQSRKRVGCFVCLIESRTCSNLTLTGSKHKISFLCD 402  
DB 14 KIYVVDLSNERAMSLTIEPRPVKQSRKRVGCFVCLIESRTCSNLTLTGSKHKISFLCD 73  
QY 403 DLTRLMNNVEKTS 416  
DB 74 DLTRLMNNVEKTS 87  
RESULT 15  
AAU69483  
ID AAU69483 standard; Protein; 167 AA.  
XX  
XX AAU69483;  
AC  
XX  
XX 29-JAN-2002 (first entry)  
DT  
XX  
XX Human purified secretory polypeptide #52.  
DE  
XX  
XX Human; purified secretory polypeptide; cell proliferative disorder;  
KW cancer; immune system disorder; neurological disorder; mental disorder;  
KW motor neuron disorder; demyelinating disorder; neuromuscular disorder;  
KW central nervous system disorder; enzyme linked immunosorbent assay;  
KW ELISA; gene therapy.  
XX  
XX  
XX Homo sapiens.  
OS  
XX  
XX MO200162918-A2.  
PN  
XX  
XX 30-AUG-2001.  
PD  
XX  
XX 01-FEB-2001; 2001WO-US03465.  
PF  
XX  
XX 24-FEB-2000; 2000US-185215P.  
PR 24-FEB-2000; 2000US-185216P.  
PR 16-MAY-2000; 2000US-205232P.  
PR 17-MAY-2000; 2000US-205286P.  
PR 17-MAY-2000; 2000US-205287P.  
PR 17-MAY-2000; 2000US-205323P.  
PR 17-MAY-2000; 2000US-205324P.  
XX  
XX (INCY-) INCYTE GENOMICS INC.  
PA

XX Panzer SR, Spiro PA, Banville SC, Shah P, Chalup MS, Chang SC;  
PI Chen A, D'sa SA, Amshey S, Dahl CR, Dam TC, Daniels SE;  
PI Dufour GE, Flores V, Fong WT, Greenwalt LB, Hillman JL, Jones AL;  
PI Liu TF, Roseberry AM, Rosen BH, Russo PD, Stockdreher TK, Daffo A;  
PI Wright RJ, Yap PE, Yu JY, Bradley DL, Bratcher SR, Chen W;  
PI Cohen HJ, Hodgson DM, Lincoln SE;  
XX  
XX WPI; 2001-648217/74.  
DR  
XX  
XX Nucleic acids encoding secretory polypeptides, useful in genetic  
PT diagnosis and therapy -  
XX  
XX  
XX Disclosure; Page 221; 237pp; English.  
XX  
XX Sequences AAU69432-AAU69511 represent purified secretory polypeptides of  
CC the invention. The polypeptides and their associated polynucleotides can  
CC be used in the treatment, prevention and diagnosis of diseases associated  
CC with inappropriate secretory protein expression. These diseases include  
CC cell proliferative disorders such as atherosclerosis and psoriasis,  
CC cancers such as leukaemia and melanoma, immune system disorders such as  
CC asthma and diabetes mellitus, neurological disorders such as epilepsy and  
CC Parkinson's disease, mental disorders such as schizophrenia and seasonal  
CC affective disorder (SAD), motor neuron disorders such as amyotrophic  
CC lateral sclerosis, demyelinating disorders such as multiple sclerosis,  
CC central nervous system disorders such as mental retardation and  
CC neurofibromatosis and neuromuscular disorders such as cerebral palsy and  
CC muscular dystrophy. Target polynucleotides in a sample can be detected by  
CC hybridising the sample with a probe sequence complementary to the target  
CC polynucleotide, under conditions in which a hybridisation complex is  
CC formed, and detecting the presence or absence of the complex. The  
CC polypeptides may also be used as antigens in the production of antibodies  
CC against secretory proteins and in assays to identify modulators of  
CC protein expression and activity. The antibodies may also be used as  
CC diagnostic agents for detecting the presence of the sequences of the  
CC invention in samples e.g. by enzyme linked immunosorbent assay (ELISA).  
XX  
XX Sequence 167 AA;  
SQ  
Query Match 4.7%; Score 39; DB 22; Length 167;  
Best Local Similarity 100.0%; Pred. No. 2e-29;  
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 107 GEVLOPSTSLPTLNRTFTMDVKAKSIGLEQPSIR 145  
DB 23 GEVLOPSTSLPTLNRTFTMDVKAKSIGLEQPSIR 61  
Search completed: February 20, 2004, 19:46:32  
Job time : 89 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2004, 19:11:57 ; Search time 47 seconds  
(without alignments)  
1710.575 Million cell updates/sec

Title: US-09-899-569a-4

Perfect score: 836  
Sequence: 1 MAGLNGVSIALLGVLLGA.....SSKDTDIPLNTQDEMPBAP 836

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 283308 seqs, 96168682 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	1.1	293	2	H70713
2	8	1.0	144	2	T50500
3	8	1.0	187	1	Q0EC30
4	8	1.0	187	2	B98089
5	8	1.0	187	2	D85934
6	8	1.0	203	2	T20750
7	8	1.0	213	2	B71607
8	8	1.0	232	2	AD2542
9	8	1.0	262	2	B83827
10	8	1.0	390	2	T41199
11	8	1.0	404	2	A35407
12	8	1.0	550	1	VGBE18
13	8	1.0	603	2	S70849
14	8	1.0	682	2	F83228
15	8	1.0	74	2	T18489
16	8	0.8	74	2	T47434
17	7	0.8	83	2	AE0713
18	7	0.8	101	2	AB3548
19	7	0.8	104	2	A33731
20	7	0.8	107	2	C83265
21	7	0.8	107	2	C86477
22	7	0.8	119	2	S08077
23	7	0.8	129	2	B90105
24	7	0.8	129	2	AC2036
25	7	0.8	132	2	S63978
26	7	0.8	132	2	B83299
27	7	0.8	140	1	F2MU10
28	7	0.8	140	2	G83318
29	7	0.8	149	2	C72611

30	7	0.8	152	2	F64441	molybdenum cofacto
31	7	0.8	154	1	UDOP07	ubiquitin / riboso
32	7	0.8	156	2	B84991	30S ribosomal prot
33	7	0.8	162	2	T17047	NADH2 dehydrogenas
34	7	0.8	162	2	T17045	NADH2 dehydrogenas
35	7	0.8	163	2	T42696	hypothetical prote
36	7	0.8	168	2	PH0139	M protein-like mol
37	7	0.8	168	2	D71266	hypothetical prote
38	7	0.8	171	2	T41924	hypothetical prote
39	7	0.8	174	2	T11517	NADH2 dehydrogenas
40	7	0.8	176	2	D97084	enzyme of dihydrof
41	7	0.8	186	2	A75281	hypothetical prote
42	7	0.8	187	2	AF3009	hypothetical prote
43	7	0.8	188	2	JC5154	K-ras protein - tu
44	7	0.8	189	2	JC4072	virulence-associat
45	7	0.8	195	2	H98274	hypothetical prote

## ALIGNMENTS

```

RESULT 1
H70713
hypothetical protein RV1509 - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 28-Jul-2000
C:Accession: H70713
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.;
; Connor, R.; Davies, R.; Devlin, K.; Fellwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A>Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295887; PMID:9634230
A:Accession: H70713
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-293 <COL>
A:Cross-references: GB:Z79701; GB:AL123456; NID:93261635; PIDN:CAB02023.1; PID:e264138; I
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: RV1509
C:Superfamily: Mycobacterium tuberculosis hypothetical protein RV1509

Query Match 1.1%; Score 9; DB 2; Length 293;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 674 VGGGVLLLS 682
DB 67 VGGGVLLLS 75

RESULT 2
T50500
hypothetical protein T22D6_90 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 08-Sep-2000
C:Accession: T50500
R:Bevan, M.; Terry, N.; Ardiles, W.; Buysaert, C.; Dasseville, R.; De Clerck, R.; De
ewes, H.W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.
submitted to the Protein Sequence Database, May 2000
A:Reference number: Z25101
A:Accession: T50500
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-144 <BEV>
A:Cross-references: EMBL:AL357612
A:Experimental source: cultivar Columbia; BAC clone T22D6
C:Genetics:
A:Map position: 5
A>Note: T22D6_90
C:Superfamily: Arabidopsis thaliana hypothetical protein T22D6_90
  
```

Query Match 1.0%; Score 8; DB 2; Length 144;  
Best Local Similarity 100.0%; Pred. No. 7.2;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 692 KKKKKKTN 699  
DB 81 KKKKKKTN 88

## RESULT 3

QOESC30  
prepilin peptidase dependent protein B precursor - Escherichia coli (strain K-12)

C:Species: Escherichia coli

C:Date: 31-Mar-1988 #sequence\_revision 05-Dec-1997 #text\_change 01-Mar-2002

C:Accession: B65065; C24137

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co

.A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; PMID:97426617; PMID:9278503

A:Accession: B65065

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-187 <BLAT>

A:Cross-references: GB:AE000366; GB:U00096; NID:g1789185; PIDN:AACT5864.1; PID:g1789189;

A:Experimental source: strain K-12, substrain MGL655

R:Finch, P.W.; Wilson, R.E.; Brown, K.; Hickson, I.D.; Tomkinson, A.E.; Emerson, P.T.

Nucleic Acids Res. 14, 4437-4451, 1986

A:Title: Complete nucleotide sequence of the Escherichia coli recC gene and of the thxA-

A:Reference number: A93625; PMID:86232583; PMID:3520484

A:Accession: C24137

A:Molecule type: DNA

A:Residues: 'MVPCFFRCQYLSMCFITGTCATLARSRNERPDTFACFLMPAQYRGMRAVSLQKNGRVVAGGFAYGKTPALSA

A:Cross-references: GB:X03966; NID:g42684; PIDN:CAA27601.1; PID:g42686

A:Genetics:

A:Gene: ppdB

A:Map position: 61 min

C:Superfamily: prepilin peptidase dependent protein B precursor

Query Match 1.0%; Score 8; DB 1; Length 187;  
Best Local Similarity 100.0%; Pred. No. 9;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 VLLIGAAR 22  
DB 22 VLLIGAAR 29

RESULT 4  
B98089  
prepilin peptidase dependent protein B [imported] - Escherichia coli (strain O157:H7, su

C:Species: Escherichia coli

C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 03-Aug-2001

C:Accession: B98089

R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.

gaasawara, N.; Yasunaga, T.; Kohara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno

A:Reference number: A99629; PMID:21156231; PMID:11258796

A:Accession: B98089

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-187 <HAY>

A:Cross-references: GB:BA000007; PIDN:BA037105.1; PID:g13363154; GSPDB:GN00154

A:Experimental source: strain O157:H7, substrain RIMD 0509952

C:Genetics:

A:Gene: ECG3682

C:Superfamily: prepilin peptidase dependent protein B precursor

QY 15 VLLIGAAR 22  
DB 22 VLLIGAAR 29

RESULT 5  
D85934

prepilin peptidase dependent protein B [imported] - Escherichia coli (strain O157:H7, su

C:Species: Escherichia coli

C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001

C:Accession: D85934

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew,

Iller, L.; Grobbeck, E.J.; Davis, N.W.; Llm, A.; Dialianta, E.; Potamoulsis, K.; Apodaca,

Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A85480; PMID:21074935; PMID:11206551

A:Accession: D85934

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-187 <STO>

A:Cross-references: GB:AE005174; NID:g12517310; PIDN:AA657936.1; GSPDB:GN00145; UWGP:Z414

A:Experimental source: strain O157:H7, substrain EDJ933

C:Genetics:

A:Gene: ppdB

C:Superfamily: prepilin peptidase dependent protein B precursor

QY 15 VLLIGAAR 22  
DB 22 VLLIGAAR 29

RESULT 6  
T20750  
hypothetical protein F11A5.4 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 11-Jan-2000

C:Accession: T20750

R:Gardner, A.

submitted to the EMBL Data Library, March 1997

A:Reference number: Z19319

A:Accession: T20750

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-203 <WIL>

A:Cross-references: EMBL:Z92830; PIDN:CAB07357.1; GSPDB:GN00023; CESP:F11A5.4

A:Experimental source: clone F11A5

C:Genetics:

A:Gene: CESP:F11A5.4

A:Map position: 5

C:Superfamily: ras transforming protein; translation elongation factor Tu homology

QY 690 CVYKXXXX 697  
DB 180 CVYKXXXX 187

RESULT 7  
B71607  
hypothetical protein PFB0740C - malaria parasite (Plasmodium falciparum)

C:Species: Plasmodium falciparum

C:Date: 13-Nov-1998 #sequence\_revision 13-Nov-1998 #text\_change 21-Jul-2000

C:Accession: B71607

R:Gardner, M.J.; Jettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.;

; Pertea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.;

Science 282, 1126-1132, 1998  
A>Title: Chromosome 2 sequence of the human malaria parasite *Plasmodium falciparum*.  
A:Reference number: A71600; MUID:99021743; PMID:9804551  
A:Accession: B71607  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-213 <GAR>  
A:Cross-references: GB:AE001415; GB:AE001362; NID:g3845264; PIDN:AACT1937.1; PID:g384526  
A:Experimental source: clone 3D7  
C:Genetics:  
A:Gene: PFB0740C

Query Match 1.0%; Score 8; DB 2; Length 213;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 692 KKKKKKTN 699  
DB 5 KKKKKKTN 12

RESULT 8  
AD2542  
hypothetical protein all7623 [imported] - *Nostoc* sp. (strain PCC 7120) plasmid pCC7120be  
C:Species: *Nostoc* sp. PCC 7120  
A>Note: *Nostoc* sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120  
C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 12-May-2003  
C:Accession: AD2542  
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Matsuabe, A.; Itiguchi, N.; Shimizu, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.  
DNA Res. 8, 205-213, 2001  
A>Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Anabaena* sp. strain PCC 7120; MUID:21595285; PMID:11759840  
A:Reference number: A18107; MUID:21595285; PMID:11759840  
A:Accession: AD2542  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-232 <KUR>  
A:Cross-references: GB:AP003602; PIDN:BA877266.1; PID:g17134708; GSPDB:GN00181  
A:Experimental source: strain PCC 7120  
C:Genetics:  
A:Gene: all7623  
A:Genome: plasmid  
C:Superfamily: bacitracin transport permease

Query Match 1.0%; Score 8; DB 2; Length 232;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 666 LTVILIAA 673  
DB 179 LTVILIAA 186

RESULT 9  
B83827  
hypothetical protein BH1418 [imported] - *Bacillus halodurans* (strain C-125)  
C:Species: *Bacillus halodurans*  
C:Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001  
C:Accession: B83827  
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira, Nucleic Acids Res. 28, 4317-4331, 2000  
A>Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and A:Reference number: A83650; MUID:20512582; PMID:11058132  
A:Accession: B83827  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-262 <STO>  
A:Cross-references: GB:AF001512; GB:BA000004; NID:g10174030; PIDN:BA805137.1; GSPDB:GN00  
A:Experimental source: strain C-125  
C:Genetics:  
A:Gene: BH1418  
C:Superfamily: hypothetical protein MJ0933

Query Match 1.0%; Score 8; DB 2; Length 262;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ALLGVLL 18  
DB 5 ALLGVLL 12

RESULT 10  
T41199  
dom34 protein homolog - fission yeast (*Schizosaccharomyces pombe*)  
C:Species: *Schizosaccharomyces pombe*  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 21-Jan-2000  
C:Accession: T41199  
R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Jimenez Martinez, J.  
submitted to the EMBL Data Library, July 1999  
A:Reference number: Z21978  
A:Accession: T41199  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-390 <MOO>  
A:Cross-references: EMBL:AL109736; PIDN:CA852153.1; GSPDB:GN00068  
A:Experimental source: strain 972h-; cosmid c1885  
C:Genetics:  
A:Gene: SPBC1885.06  
A:Map position: 3  
A:Introns: 27/2; 45/2; 145/1; 180/3  
C:Superfamily: cell division protein MJ0174

Query Match 1.0%; Score 8; DB 2; Length 390;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 647 SGKQIDL 654  
DB 352 SGKQIDL 359

RESULT 11  
A35407  
tryptophan synthase (EC 4.2.1.20) beta chain - *Thermus aquaticus*  
C:Species: *Thermus aquaticus*  
C:Date: 14-Sep-1990 #sequence\_revision 14-Sep-1990 #text\_change 16-Jul-1999  
C:Accession: A35407  
R:Koyama, Y.; Furukawa, K.  
J. Bacteriol. 172, 3490-3495, 1990  
A>Title: Cloning and sequence analysis of tryptophan synthetase genes of an extreme thermophile *T. thermophilus* cells.  
A:Reference number: A35407; MUID:90264352; PMID:2188962  
A:Accession: A35407  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-404 <KOR>  
A:Cross-references: GB:M32108; NID:g155132; PIDN:AAA27508.1; PID:g155133  
A:Experimental source: strain HB27  
C:Genetics:  
A:Gene: trpB  
C:Function: catalyzes conversion of indoleglycerol phosphate and serine to tryptophan  
A:Pathway: tryptophan biosynthesis  
A>Note: cofactor pyridoxal phosphate; last step in pathway  
C:Superfamily: tryptophan synthase beta chain; tryptophan synthase beta chain homology  
C:Keywords: carbon-oxygen lyase; hydro-lyase; phosphoprotein; pyridoxal phosphate; tryptophan synthase; tryptophan synthase beta chain homology <TRPS>  
F:12-397/Domain: tryptophan synthase beta chain homology <TRPS>  
F:94/Active site: His #status predicted  
F:95/Binding site: pyridoxal phosphate (lys) (covalent) #status predicted

Query Match 1.0%; Score 8; DB 2; Length 404;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 670 LIAAVGG 677

Db 235 LIAAVGGG 242

## RESULT 12

VGBE18

glycoprotein E - human herpesvirus 1

N:Alternate names: US8

C:Species: human herpesvirus 1

C&gt;Date: 17-Mar-1987 #sequence\_revision 17-Mar-1987 #text\_change 16-Jul-1999

C:Accession: A03733; A45696

R:McGeoch, D.J.; Dolan, A.; Donald, S.; Rixon, F.J.

J. Mol. Biol. 161, 1-13, 1985

A&gt;Title: Sequence determination and genetic content of the short unique region in the ge

A:Reference number: A00656; MUID:85160822; PMID:2984429

A:Accession: A03733

A:Molecule type: DNA

A:Residues: 1-550 &lt;MCG&gt;

A:Cross-references: GB:X02138, NID:959865; PIDN:CAA26062.1; PID:959882

A:Experimental source: strain 17

R:Georgopoulou, U.; Michaelidou, A.; Roizman, B.; Mavromara-Nazos, P.

J. Virol. 67, 3961-3968, 1993

A&gt;Title: Identification of a new transcriptional unit that yields a gene product within

A:Reference number: A45696; MUID:93287213; PMID:8389914

A:Accession: A45696

A:Molecule type: DNA

A:Residues: 438-550 &lt;GEO&gt;

A:Cross-references: GB:S62895; NID:9386127; PIDN:AA27080.1; PID:9386128

A:Experimental source: R35

A&gt;Note: sequence extracted from NCBI backbone (NCBIN:133646, NCBIPI:133647)

C:Superfamily: herpesvirus glycoprotein E

C:Keywords: glycoprotein

F:124,243,501/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 1.0%; Score 8; DB 1; Length 550;

Best Local Similarity 100.0%; Pred. No. 23;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 679 LLLSALGL 686

428 LLLSALGL 435

## RESULT 13

S70849

cholinesterase (EC 3.1.1.8) - mouse

N:Alternate names: butyrylcholine esterase

C:Species: Mus musculus (house mouse)

C&gt;Date: 28-Oct-1996 #sequence\_revision 08-Nov-1996 #text\_change 18-Jun-1999

C:Accession: S70849; S15680; A39768

R:Taylor, P.

submitted to the EMBL Data Library, August 1992

A:Reference number: S70849

A:Accession: S70849

A:Molecule type: nucleic acid

A:Residues: 1-603 &lt;TAY&gt;

A:Cross-references: EMBL:M99492; NID:9191579; PIDN:AAA37328.1; PID:9191580

R:Rachinsky, T.L.; Camp, S.; Li, Y.; Ekstrom, T.J.; Newton, M.; Taylor, P.

Neuron 5, 317-327, 1990

A&gt;Title: Molecular cloning of mouse acetylcholinesterase: tissue distribution of alterna

A:Reference number: JH0314; MUID:90380429; PMID:2400605

A:Accession: S15680

A&gt;Status: nucleic acid sequence not shown

A:Molecule type: nucleic acid

A:Residues: 30-128, P, 130-603 &lt;RAC&gt;

A:Cross-references: EMBL:M99492

R:Arpagaus, M.; Chatonnet, A.; Masson, P.; Newton, M.; Vaughan, T.A.; Bartels, C.F.; Nog

J. Biol. Chem. 266, 6966-6974, 1991

A&gt;Title: Use of the polymerase chain reaction for homology probing of butyrylcholinester

A:Reference number: A39768; MUID:91201348; PMID:2016308

A:Accession: A39768

A&gt;Status: preliminary

A:Molecule type: DNA  
A:Residues: 97-128, P, 130-237 <ARP>  
A:Superfamily: Cholinesterase; cholinesterase homology  
C:Keywords: carboxylic ester hydrolase; glycoprotein  
F:57-557/Domain: cholinesterase homology <CHE>

Query Match 1.0%; Score 8; DB 2; Length 603;

Best Local Similarity 100.0%; Pred. No. 25;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 613 TRAEIRFS 620

486 TRAEIRFS 493

## RESULT 14

hypothetical protein PA3340 [imported] - Pseudomonas aeruginosa (strain PA01)

C:Species: Pseudomonas aeruginosa

C&gt;Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000

C:Accession: F83228

R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Bri

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lapidig, K.; Llm,

.; Loly, S.; Olson, M.V.

Nature 406, 959-964, 2000

A&gt;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathog

A:Reference number: A82950; MUID:20437337; PMID:10984043

A:Accession: F83228

A&gt;Status: preliminary

A:Molecule type: DNA

A:Residues: 1-682 &lt;STO&gt;

A:Cross-references: GB:AE004756; GB:AE004091; NID:99949466; PIDN:AA606728.1; GSPDB:GN0013

A:Experimental source: strain PA01

C:Genetics:

A:Gene: PA3340

Query Match 1.0%; Score 8; DB 2; Length 682;

Best Local Similarity 100.0%; Pred. No. 28;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 11 ALLGVLL 18

364 ALLGVLL 371

## RESULT 15

T18489

hypothetical protein C0820w - malaria parasite (Plasmodium falciparum)

C:Species: Plasmodium falciparum

C&gt;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jun-2000

C:Accession: T18489

R:Lawson, D.; Bowman, S.; Barrell, B.

submitted to the EMBL Data Library, August 1997

A:Reference number: Z18935

A:Accession: T18489

A&gt;Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-4981 &lt;LAW&gt;

A:Cross-references: EMBL:Z98551; NID:e1331903; PID:e1331910; PIDN:CA011128.1

C:Genetics:

A:Map position: 3

A&gt;Note: C0820w

Query Match 1.0%; Score 8; DB 2; Length 4981;

Best Local Similarity 100.0%; Pred. No. 1.6e+02;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 692 KKKKKKTN 699

1439 KKKKKKTN 1446

Search completed: February 20, 2004, 19:50:19



Mon Feb 23 08:34:08 2004

us-09-899-569a-4.011.rpr

Page 5

Job time : 49 secs

---

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: February 20, 2004, 13:55:12 ; Search time 27 Seconds

(without alignments)  
1456.086 Million cell updates/sec

Title: US-09-899-569a-4

Perfect score: 836  
Sequence: 1 MAGLNCVSLALGVLLGA.....SSKQDIPLNTQSEMPAP 836

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 127863 seqs, 47026705 residues

Word size: 0

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: SwissProt\_41.1\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9	1.1	173	PSAL_MASLA	O31126 mastigoclad
2	9	1.1	293	YF09_MYCTU	P71788 mycobacteri
3	8	1.0	187	PP08_ECOLI	P08371 escherichia
4	8	1.0	404	TRPB_THEH	P16609 thermus the
5	8	1.0	485	ZDH1_HUMAN	O8wtx9 homo sapien
6	8	1.0	485	ZDH1_MOUSE	O8rc09 mus musculu
7	8	1.0	550	USHA_SALPU	O9fn37 salmoneilla
8	8	1.0	550	VGLE_HSV1	P04488 herpes simp
9	8	1.0	603	CHLE_MOUSE	O03311 mus musculu
10	8	0.8	78	R27A_DICDI	P14797 dictyosteli
11	8	0.8	104	SLIB_RAT	P09916 ratius norv
12	7	0.8	132	SNPA_STRCS	P56406 streptomyce
13	7	0.8	140	PSBR_ARATH	P27202 arabidopsis
14	7	0.8	141	PSBR_BRACH	P49108 brachista ca
15	7	0.8	152	MOAC_METVA	O58535 methanococc
16	7	0.8	152	PSAL_GUITR	O74699 guillardi
17	7	0.8	156	RS7_BUCAL	P57594 buchnera ap
18	7	0.8	156	RS7_BUCAL	P57594 buchnera ap
19	7	0.8	174	NU6M_PAPHA	O9xxxs papio hamad
20	7	0.8	188	RASK_MELDA	P79800 melagris g
21	7	0.8	216	SSPN_MOUSE	O62147 mus musculu
22	7	0.8	219	ATPD_CHLRE	O42687 chlamydomon
23	7	0.8	219	YIIL_STRCO	O86576 streptomyce
24	7	0.8	227	COX2_SITGR	P29879 sitophilus
25	7	0.8	243	SSPN_HUMAN	O14714 homo sapien
26	7	0.8	265	BXB2_HAETN	P19391 haemophilus
27	7	0.8	265	BXB2_HAETN	P22235 haemophilus
28	7	0.8	278	RCEL_CHRYI	P51762 chromatiu
29	7	0.8	281	YE94_XYLFA	O9pd85 xyloella fas
30	7	0.8	285	IF2B_YEAST	P08064 saccharomyc
31	7	0.8	300	ERA_MYCLB	O49768 mycobacteri
32	7	0.8	312	IF2B_DROME	P41375 drosophila
33	7	0.8	328	P111_HUMAN	O9bks5 homo sapien

34	7	0.8	331	IF2B_MOUSE	O99145 mus musculu
35	7	0.8	332	P111_MOUSE	O9cex5 mus musculu
36	7	0.8	333	IF2B_HUMAN	P20042 homo sapien
37	7	0.8	333	IF2B_RABIT	P41035 oryctolagus
38	7	0.8	335	BTUC_YERPE	O8zdx4 yerinia pe
39	7	0.8	335	LEP4_MYXXA	O30387 myxococcus
40	7	0.8	343	YD57_METVA	O58752 methanococc
41	7	0.8	346	YOC2_CAEEL	O09261 caenorhabdi
42	7	0.8	371	RLX1_SALTY	P14492 salmoneilla
43	7	0.8	380	CYB_THOIS	O9xw9 thomascomye
44	7	0.8	381	CYB_AULMI	O9xnv4 auliscomys
45	7	0.8	381	CYB_ELIMO	O9xnv3 eliomodonti

## ALIGNMENTS

RESULT 1  
PSAL\_MASLA STANDARD; PRT; 173 AA.  
ID PSAL\_MASLA  
AC O31126;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Photosystem I reaction center subunit XI (PSI-X) (PSI subunit V).  
GN PSAL.  
OS Mastigocladus laminosus (Fischerella sp.).  
OC Bacteria; Cyanobacteria; Stigonematales; Mastigocladus.  
OX NCBI\_Taxid=83541;  
RN (1)  
RP SEQUENCE FROM N.A.  
RC STRAIN=PCC.7605;  
RA He Z.-Y., Chitnis P.R., Nechushtal R.;  
RT "Molecular cloning of the psal gene for photosystem I subunit XI from  
the thermophilic cyanobacterium Mastigocladus laminosus.";  
RL (In) Plant Gene Register PGR98-025.  
CC -I- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. CELLULAR  
THYLAKOID MEMBRANE (PROBABLE).  
CC -I- SIMILARITY: BELONGS TO THE PSAL FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: AF030003; AAC04841.1; -  
DR HAMAP: MF\_00447; -; 1.  
DR InterPro: IPR003757; PSI\_Psal.  
DR Pfam: PF02605; Psal.1.  
DR ProDom: PD005947; PSI\_Psal.1.  
KW Photosystem I; Photosynthesis; Transmembrane.  
FT TRANSMEM 92 112  
FT TRANSMEM 142 162  
SQ SEQUENCE 173 AA; 18444 MW; 24E04D10B908715D CRC64;  
Query Match 1.1%; Score 9; DB 1; Length 173;  
Best Local Similarity 100.0%; Pred. No. 0.34;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
CY 680 LLSALGLII 688  
DB 95 LLSALGLII 103  
RESULT 2  
ID YF09\_MYCTU STANDARD; PRT; 293 AA.  
AC P71788;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)

```

DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein RV1509.
GN RV1509 OR MT1557 OR MTCV277.31.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OK NCBI_TaxId=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Broesch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Baaham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Fellwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornby T., Javels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Esmolaeva M.D., Salzberg S.L.,
RA Delcher A., Uterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z79701; CAB02023.1; -.
DR EMBL: AE007023; AAK45824.1; -.
DR PIR: H70713; H70713.
DR TIGR: MT1557; -.
DR Tuberculist; RV1509; -.
KM Hypothetical protein; Complete proteome.
SQ SEQUENCE 293 AA; 33055 MW; 7393B5725FE71EB7 CRC64;

Query Match 1.1%; Score 9; DB 1; Length 293;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 674 VGGGVILLLS 682
DB 67 VGGGVILLLS 75

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE=86232583; PubMed=3520484;
RA Finch P.W., Wilson R.E., Brown K., Hickson I.D., Tomkinson A.E.,
RA Emerson P.T.;
RT "Complete nucleotide sequence of the Escherichia coli recC gene and
RT of the thvA-recC intergenic region."
RL Nucleic Acids Res. 14:4437-4451(1986).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
RN [3]
RP IDENTIFICATION.
RX MEDLINE=95020523; PubMed=7934814;
RA Hobbs M., Matlick J.S.;
RT "Common components in the assembly of type 4 fimbriae, DNA transfer
RT systems, filamentous phage and protein-secretion apparatus: a general
RT system for the formation of surface-associated protein complexes."
RL Mol. Microbiol. 10:233-243(1993).
RN [4]
RP FUNCTION: NOT YET KNOWN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X03966; CAA27601.1; ALT_INIT.
DR EMBL: U29581; AAB40472.1; -.
DR EMBL: AE000366; AAC75864.1; -.
DR PIR: B65065; Q0EC30.
DR Ecogene; EG11156; ppDB.
DR InterPro; IPR001120; Prok N methyltn.
DR PROSITE; PS00409; PROKAR_NTER_METHYL. 1.
KM Methylation; Complete proteome.
FT PROPEP 1 7
FT CHAIN 8 187
FT MOD RES 8 8 METHYLATION (BY SIMILARITY).
SQ SEQUENCE 187 AA; 20519 MW; 433303D846D786F0 CRC64;

Query Match 1.0%; Score 8; DB 1; Length 187;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 VLLLGAR 22
DB 22 VLLLGAR 29

```

```

RESULT 4
ID TRPB_THETH STANDARD; PRT; 404 AA.
AC P16609;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Tryptophan synthase beta chain (EC 4.2.1.20).
GN TRPB.
OS Thermus thermophilus.
OC Bacteria; Deinococcus-Thermus; Thermococci; Thermales; Thermaceae;
OC Thermus.
OK NCBI_TaxId=274;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HB27;

```

```

RX MEDLINE=90264352; PubMed=2188962;
RA Koyama Y., Furukawa K.;
RT "Cloning and sequence analysis of tryptophan synthetase genes of an
RT extreme thermophile, Thermus thermophilus H27: plasmid transfer from
RT replica-plated Escherichia coli recombinant colonies to competent T.
RT thermophilus cells";
RL J. Bacteriol. 172:3490-3495(1990).
RN [2]
RP SEQUENCE OF 1-45 FROM N.A.
RC STRAIN=HB27;
RX MEDLINE=91130853; PubMed=2283046;
RA Koyama Y., Arikawa Y., Furukawa K.;
RT "A plasmid vector for an extreme thermophile, Thermus thermophilus";
RL FEMS Microbiol. Lett. 60:97-101(1990).
CC -1- FUNCTION: The beta subunit is responsible for the synthesis of L-
CC tryptophan from indole and L-serine.
CC -1- CATALYTIC ACTIVITY: L-serine + L-(indol-3-yl)glycerol 3-phosphate
CC = L-tryptophan + glyceraldehyde 3-phosphate.
CC -1- COFACTOR: Pyridoxal phosphate (By similarity).
CC -1- PATHWAY: Tryptophan biosynthesis; fitch (last) step.
CC -1- SUBUNIT: Tetramer of two alpha and two beta chains (By
CC similarity).
CC -1- SIMILARITY: BELONGS TO THE TRPB FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M32108; AAA27508.1; -.
DR EMBL; X58673; CAA41527.1; -.
DR HSSP; P00933; 2MSY.
DR HAMAP; MF_00133; -.
DR InterPro; IPR001926; B6_enzyme_beta.
DR InterPro; IPR006653; Trp_synth_b_rel.
DR InterPro; IPR006654; Trp_synth_beta.
DR Pfam; PF00291; PALP; 1.
DR TIGRfam; TIGR00263; trpb; 1.
DR PROSITE; PS00168; TRP_SYNTHASE_BETA; 1.
DR Tryptophan biosynthesis; PYRIDOXAL PHOSPHATE (BY SIMILARITY).
FT BINDING 95
FT SEQUENCE 404 AA; 43809 MW; 74D09A1EEC2A0466 CRC64;
SQ
Query Match 1.0%; Score 8; DB 1; Length 404;
Best Local Similarity 100.0%; Pred. No. 7.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 670 LIAAVGGG 677
DB 235 LIAAVGGG 242

```

```

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Ditchenko L., Marzina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
RA Brownstein W.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.J.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Boulford G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalski M., Smallus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE OF 1-293 FROM N.A., AND TISSUE SPECIFICITY.
RC TISSUE=PANCREAS;
RX MEDLINE=99321009; PubMed=10395086;
RA Puttina T., Wong P., Gentileman S.;
RT "The DHHC domain: a new highly conserved cysteine-rich motif.";
RL Mol. Cell. Biochem. 195:219-226(1999).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- TISSUE SPECIFICITY: Expressed at high levels in fetal lung, kidney
CC and heart. Expressed at lower levels in adult pancreas and lung.
CC -1- SIMILARITY: Contains 1 DHHC-type zinc finger.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; BC021908; AAH21908.1; -.
DR EMBL; U90653; AAB86591.2; -.
DR Genew; HGNC:17916; ZDHHC1.
DR InterPro; IPR001594; Znf_DHHC.
DR Pfam; PF01529; ZF-DHHC; 1.
DR ProDom; PD003041; Znf_DHHC; 1.
DR PROSITE; PS50216; ZF-DHHC; 1.
DR Transmembrane; Znf-finger.
FT TRANSMEM 53
FT TRANSMEM 73
FT TRANSMEM 78
FT TRANSMEM 98
FT TRANSMEM 186
FT TRANSMEM 206
FT TRANSMEM 242
FT ZN_FING 134
FT ZN_FING 184
FT SEQUENCE 485 AA; 54818 MW; 6B75B07D7D82F358 CRC64;
SQ
Query Match 1.0%; Score 8; DB 1; Length 485;
Best Local Similarity 100.0%; Pred. No. 8.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 11 ALLGVLL 18
DB 186 ALLGVLL 193

```

```

RESULT 5
ZDHL_HUMAN STANDARD; PRT; 485 AA.
ID ZDHL_HUMAN STANDARD; PRT; 485 AA.
AC Q8WXT9; O15461;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
DE Zinc finger DHHC domain containing protein 1 (Zinc finger protein 377)
DE (DHHC-domain-containing cysteine-rich protein 1).
GN ZDHHC1 OR ZNF377 OR C16ORF1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
CX NCI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A.
RP TISSUE=uterus;
RX MEDLINE=22388257; PubMed=12477932;

```

```

RESULT 6
ZDHL_MOUSE STANDARD; PRT; 485 AA.
ID ZDHL_MOUSE STANDARD; PRT; 485 AA.
AC Q8RON9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
DE Zinc finger DHHC domain containing protein 1 (Fragment).
GN ZDHHC1.

```

OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Colon;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ussdt T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mollay S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gnarthe P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahney J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RP (2)  
 RT TISSUE SPECIFICITY.  
 RC TISSUE=Pancreas;  
 RX MEDLINE=99321009; PubMed=10395086;  
 RA Pucillina T., Wong P., Gentleman S.;  
 RT "The DHHC domain: a new highly conserved cysteine-rich motif.";  
 RL Mol. Cell. Biochem. 195:219-226(1999).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -1- TISSUE SPECIFICITY: Expressed at high levels in fetal lung and  
 CC heart. Expressed at lower levels in fetal liver and brain. Also  
 CC detected in adult islet cells of pancreas, Leydig cells of testis,  
 CC retina and molecular layer of cerebellum.  
 CC -1- SIMILARITY: Contains 1 DHHC-type zinc finger.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: BC026570; AAH26570.1; -;  
 DR InterPro: IPR001594; Znf.DHHC.  
 DR Pfam: PF01529; Zf-DHHC.1.  
 DR ProDom: PD003041; Znf.DHHC.1.  
 DR PROSITE: PS00216; ZF-DHHC; 1.  
 KW Transmembrane; Zinc-finger.  
 FT FT NON TER 1 1  
 FT TRANSMEM 51 71 POTENTIAL.  
 FT TRANSMEM 76 96 POTENTIAL.  
 FT TRANSMEM 184 204 POTENTIAL.  
 FT TRANSMEM 240 260 POTENTIAL.  
 FT ZN FING 132 182 DHHC-TYPE.  
 SO SEQUENCE 485 AA; 53107 MW; 8668652C90A49139 CRC64;

Query Match 1.0%; Score 8; DB 1; Length 485;  
 Best Local Similarity 100.0%; Pred. No. 8.9;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 ALLGVLL 18  
 DB 184 ALLGVLL 191

RESULT 7  
 ID USHA\_SALPU STANDARD; PRT; 550 AA.  
 AC 09RN37;  
 DT 15-SEP-2003 (Rel. 42, Created)  
 DT 15-SEP-2003 (Rel. 42, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Protein usha precursor [includes: UDP-sugar hydrolase (EC 3.6.1.45)  
 DE (UDP-sugar diphosphatase) (UDP-sugar pyrophosphatase); 5'-nucleotidase  
 DE (EC 3.1.3.5) (5'-NT)].  
 GN USHA.  
 OS Salmonella pullorum.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Salmonella.  
 OX NCBI\_TaxID=605;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21322714; PubMed=11429465;  
 RA Innes D., Beacham I.R., Beven C.-A., Douglas M., Laird M.W.,  
 RA Joly J.C., Burns D.M.;  
 RT "The cryptic usha gene (ushA(c)) in natural isolates of Salmonella  
 RT enterica (serotype Typhimurium) has been inactivated by a single  
 RT missense mutation.";  
 RL Microbiology 147:1867-1896(2001).  
 CC -1- FUNCTION: DEGRADATION OF EXTERNAL UDP-GLUCOSE TO URIDINE  
 CC MONOPHOSPHATE AND GLUCOSE-1-PHOSPHATE, WHICH CAN THEN BE USED BY  
 CC THE CELL (By similarity).  
 CC -1- CATALYTIC ACTIVITY: UDP-sugar + H(2)O = UMP + sugar 1-phosphate.  
 CC -1- CATALYTIC ACTIVITY: A 5'-ribonucleotide + H(2)O = a ribonucleoside  
 CC + phosphate.  
 CC -1- COFACTOR: BINDS TWO ZINC IONS (By similarity).  
 CC -1- SUBUNIT: Monomer (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Periplasmic (By similarity).  
 CC -1- SIMILARITY: BELONGS TO THE 5'-NUCLEOTIDASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AF188727; AAF05581.1; -;  
 DR HSSP: P07024; IUSH.  
 DR InterPro: IPR006179; 5\_nucleotidase.  
 DR InterPro: IPR006146; 5\_NUCLEOTIDASE\_2.  
 DR Pfam: PF02872; 5\_nucleotidasec1.  
 DR Pfam: PF00149; Metallophos; 1.  
 DR PRINTS: PR01607; APYRASEFAMLY.  
 DR PROSITE: PS00785; 5\_NUCLEOTIDASE\_1; 1.  
 DR PROSITE: PS00786; 5\_NUCLEOTIDASE\_2; 1.  
 KW Hydrolase; Zinc; Periplasmic; Signal; Multifunctional enzyme.  
 FT FT SIGNAL 1 25  
 FT CHAIN 26 550  
 FT ACT SITE 117 117  
 FT ACT SITE 120 120  
 FT METAL 41 41  
 FT METAL 43 43  
 FT METAL 84 84  
 FT METAL 116 116  
 FT METAL 217 217  
 FT METAL 252 252  
 FT METAL 254 254  
 SO SEQUENCE 550 AA; 60604 MW; 655BC78F0D4F62F CRC64;

Query Match 1.0%; Score 8; DB 1; Length 550;  
 Best Local Similarity 100.0%; Pred. No. 9.9;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 675 GGGVLLS 682  
 |||||

Db 74 GCGVLLS 81

RESULT 8  
VGLV\_HSV11 STANDARD; PRT; 550 AA.

AC P04468; 13-AUG-1987 (Rel. 05, Created)  
DT 13-AUG-1987 (Rel. 05, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Glycoprotein E precursor.  
GN GE OR US8  
OS Herpes simplex virus (type 1 / strain 17).  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Alphaherpesvirinae; Simplexvirus.  
OX NCBI\_TaxID=10299;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA McGeoch D.J., Dolan A., Donald S., Rixon F.J.;  
RT "Sequence determination and genetic content of the short unique  
region in the genome of herpes simplex virus type 1."  
J. Mol. Biol. 181:1-13(1985).  
CC -1- MISCELLANEOUS: THERE ARE SEVEN EXTERNAL GLYCOPROTEINS IN HSV1 AND  
2: GH, GB, GC, GD, GI, AND GE.  
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN E FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X14112; CAA32272.1; -  
DR EMBL; X02138; CAA26062.1; -  
DR EMBL; L00036; AAA96680.1; -  
DR PIR; A03733; VGBE18.  
DR InterPro; IPR003404; Herpes\_glycoprotein.  
DR Pfam; PF02480; Herpes\_GE\_1.  
KM Glycoprotein; Signal; Transmembrane.  
FT SIGNAL 1 21  
FT CHAIN 1 21  
FT CARBOHYD 22 550  
FT CARBOHYD 124 243  
FT CARBOHYD 501 501  
FT SEQUENCE 550 AA; 59093 MW; B88271E5B7E34776 CRC64;  
SQ

Query Match 1.0%; Score 8; DB 1; Length 550;  
Best Local Similarity 100.0%; Pred. No. 9.9;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 679 LLLSALGL 686  
Db 428 LLLSALGL 435

RESULT 9  
CHLE\_MOUSE STANDARD; PRT; 603 AA.

AC Q03311; 01-OCT-1993 (Rel. 27, Created)  
DT 01-OCT-1993 (Rel. 27, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE Cholinesterase precursor (EC 3.1.1.8) (acylcholine acylhydrolase)  
DE (Choline esterase II) (Butyrylcholine esterase)  
DE (Pseudocholinesterase).  
GN BCHE.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;

RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90380429; PubMed=2400605;  
RA Rachinsky T.L., Camp S., Li Y., Ekstrom T.J., Newton M., Taylor P.;  
RT "Molecular cloning of mouse acetylcholinesterase: tissue distribution  
of alternatively spliced mRNA species."  
Neuron 5:317-327(1990).  
RN [2]  
RP SEQUENCE OF 97-237 FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=91201348; PubMed=2016308;  
RA Arpegaus M., Chatomiet A., Maason P., Newton M., Vaughan T.A.,  
RA Bartels C.F., Nogueira C.P., La Du B.N., Lockridge O.;  
RT "Use of the polymerase chain reaction for homology probing of  
butyrylcholinesterase from several vertebrates."  
J. Biol. Chem. 266:6966-6974(1991).  
CC -1- CATALYTIC ACTIVITY: An acylcholine + H(2)O = choline + a  
carboxylic acid anion.  
CC -1- SUBUNIT: HOMOTETRAMER. THE TETRAMER IS COMPOSED OF TWO DIMERS. THE  
TWO SUBUNITS IN A DIMER ARE LINKED BY A DISULFIDE BOND.  
CC -1- TISSUE SPECIFICITY: PRESENT IN MOST CELLS (EXCEPT ERYTHROCYTES).  
CC -1- MISCELLANEOUS: CHOLINESTERASE IS HIGHLY REACTIVE WITH  
ORGANOPHOSPHATE ESTERS.  
CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/lipase family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M99492; AAA37328.1; -  
DR PIR; S70849; S70849.  
DR HSP; S70849; S70849.  
DR GSD; MGI:894278; Bche.  
DR InterPro; IPR002018; Carboxylesterase.  
DR InterPro; IPR000997; Cholinesterase.  
DR InterPro; IPR000379; Ser\_estra\_site.  
DR Pfam; PF00135; Coesterase\_1.  
DR PRINTS; PR00878; CHOLINESTERASE.  
DR PROSITE; PS00122; CARBOXYLESTERASE\_B\_1; 1.  
DR PROSITE; PS00941; CARBOXYLESTERASE\_B\_2; 1.  
KM Hydrolase; Serine esterase; Glycoprotein; Signal.  
FT SIGNAL 1 29  
FT CHAIN 1 29  
FT ACT\_SITE 227 227  
FT ACT\_SITE 354 354  
FT ACT\_SITE 467 467  
FT DISULFID 94 121  
FT DISULFID 281 292  
FT DISULFID 429 548  
FT DISULFID 600 600  
FT CARBOHYD 86 86  
FT CARBOHYD 135 135  
FT CARBOHYD 270 270  
FT CARBOHYD 370 370  
FT CARBOHYD 484 484  
FT CARBOHYD 510 510  
FT CARBOHYD 515 515  
FT CONFLICT 129 129  
FT SEQUENCE 603 AA; 68521 MW; 719B1B220D1E5367 CRC64;  
SQ

Query Match 1.0%; Score 8; DB 1; Length 603;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 613 TRAEIIPS 620  
Db 486 TRAEIIPS 493

```

RESULT 10
R27A.DICDI STANDARD; PRT; 78 AA.
ID SLIB_RAT
AC P14797;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE 40S ribosomal protein S27a.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
OX NCBI_TaxId=44689;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89352609; PubMed=2548604;
RA Ohmachi T., Giorda R., Shaw D.R., Ennis H.L.;
RT "Molecular organization of developmentally regulated Dictyostelium
discoidium ubiquitin cDNAs."
RL Biochemistry 28:5226-5231(1989).
CC -1- MISCELLANEOUS: THIS RIBOSOMAL PROTEIN IS SYNTHESIZED AS A
C-TERMINAL EXTENSION PROTEIN (CEP) OF UBIQUITIN.
CC -1- SIMILARITY: BELONGS TO THE S27AE FAMILY OF RIBOSOMAL PROTEINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M23750; AAA33264.1; ALT_INIT.
DR DickeyDb; DD01026;
DR InterPro; IPR002906; Ribosomal_S27.
DR Pfam; PF01599; Ribosomal_S27.1.
KW Ribosomal protein; zinc-finger; Metal-binding.
FT DOMAIN 1 24 LYS-RICH (HIGHLY BASIC).
FT ZN_FING 46 69 C4-TYPE.
SQ SEQUENCE 78 AA; 8915 MW; F77B3E09D1B7246F CRC64;

Query Match 0.8%; Score 7; DB 1; Length 78;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 692 KKKKKKT 698
Db 3 KKKKKKT 9

RESULT 11
SLIB_RAT STANDARD; PRT; 104 AA.
ID SLIB_RAT
AC P09916;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Somatoliberin precursor (Growth hormone-releasing factor) (GRF)
DE (Growth hormone-releasing hormone) (GHRH).
GN GHRH.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85163768; PubMed=3920534;
RA Mayo K.E., Cerelli G.M., Rosenfeld M.G., Evans R.M.;
RT "Characterization of cDNA and genomic clones encoding the precursor
to rat hypothalamic growth hormone-releasing factor."
RL Nature 314:464-467(1985).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92020929; PubMed=1924334;
RA Gonzalez-Crespo S., Boronati A.;

```

```

RT "Expression of the rat growth hormone-releasing hormone gene in
RT placenta is directed by an alternative promoter."
RT Proc. Natl. Acad. Sci. U.S.A. 88:8749-8753(1991).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Testis;
RX MEDLINE=95203210; PubMed=7895659;
RA Sivasubava C.H., Montes B.S., Rothrock J.K., Peredo M.J.,
RA Pescovitz O.H.;
RT "Presence of a spermatogenic-specific promoter in the rat growth
RT hormone-releasing hormone gene."
RL Endocrinology 136:1502-1508(1995).
RN [4]
RP SEQUENCE OF 31-73.
RC TISSUE=Hypothalamus;
RX MEDLINE=83219259; PubMed=6406907;
RA Spies J., Rivier J., Vale W.;
RT "Characterization of rat hypothalamic growth hormone-releasing
RT factor."
RL Nature 303:532-535(1983).
CC -1- FUNCTION: GRF IS RELEASED BY THE HYPOTHALAMUS AND ACTS ON THE
CC ADENOHYPOPHYSIS TO STIMULATE THE SECRETION OF GROWTH HORMONE.
CC -1- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X02319; NOT_ANNOTATED_CDS.
DR EMBL: X02335; CAA26194.1; JOINED.
DR EMBL: X02320; CAA26194.1; JOINED.
DR EMBL: X02321; CAA26194.1; JOINED.
DR EMBL: X02322; CAA26194.1; JOINED.
DR EMBL: W73486; AAA41220.1; -.
DR EMBL: U10156; AAC52184.1; -.
DR PIR; A32731; A32731.
DR InterPro; IPR000532; Glucagon.
DR Pfam; PF00123; hormone2; 1.
DR SMART; SM00070; GLUCN.1.
DR PROSITE; PS00260; GLUCAGON.1.
KW Glucagon family; Signal; Hypothalamus.
FT SIGNAL 1 19
FT PEPTIDE 31 73 SOMATOLIBERIN.
SQ SEQUENCE 104 AA; 12266 MW; F9C17485742B2887 CRC64;

Query Match 0.8%; Score 7; DB 1; Length 104;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 387 LTLTSGS 393
Db 10 LTLTSGS 16

RESULT 12
SNPA_STRCS STANDARD; PRT; 132 AA.
ID SNPA_STRCS
AC P56406;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Extracellular small neutral protease (EC 3.4.24.77) (Snappalysin)
DE (SCNP).
GN SCNP.
OS Streptomyces caespitosus.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxId=53502;
RN [1]

```



RP SEQUENCE.  
 RA MEDLINE=96067714; PubMed=7588817;  
 RX Harada S., Kinoshita T., Kasai N., Tsunawawa S., Sakiyama F.;  
 RT "Complete amino acid sequence of a zinc metalloendoprotease from  
 RT Streptomyces caespitosus";  
 RL Eur. J. Biochem. 233:683-686(1995).  
 RN [2]  
 RP X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).  
 RX MEDLINE=97244628; PubMed=9089404;  
 RA Kurisu G., Kinoshita T., Sugimoto A., Nagara A., Kai Y., Kasai N.,  
 RA Harada S.;  
 RT "Structure of the zinc endoprotease from Streptomyces caespitosus";  
 RL J. Biochem. 121:304-308(1997).  
 CC -1- FUNCTION: SPECIFICALLY HYDROLYZES THE PEPTIDE BOND AT THE IMINO  
 CC SIDE OF AROMATIC RESIDUES.  
 CC -1- CATALYTIC ACTIVITY: Hydrolyzes proteins with a preference for Tyr  
 CC or Phe in the P1' position. Has no action on amino-acid p-  
 CC nitroanilides.  
 CC -1- COFACTOR: BINDS 1 ZINC ION.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M7.  
 DR PIR: S63978; S63978.  
 DR PDB: 1KUH; 12-MAR-97.  
 DR PDB: 1C7K; 25-MAR-01.  
 DR MEROPS: M07.001; -;  
 DR InterPro: IPR000013; Peptidase\_M7.  
 DR InterPro: IPR006025; Zn\_MTPeptidase.  
 DR Pfam: PF02031; Peptidase\_M7.1.  
 DR PRINTS: PR00787; NEUTRALPEPTASE.  
 DR ProDom: PD016028; Peptidase\_M7.1.  
 DR PROSITE: PS00142; ZINC\_PROTASE; FALSE\_NEG.  
 KW Hydrolyase; Metalloprotease; Zinc; 3D-structure.  
 FT METAL 83 83 ZINC (CATALYTIC).  
 FT ACT SITE 84 84  
 FT METAL 87 87 ZINC (CATALYTIC).  
 FT METAL 93 93 ZINC (CATALYTIC).  
 FT DISULFID 112 112  
 FT STRAND 2 10  
 FT HELIX 12 14  
 FT HELIX 15 26  
 FT STRAND 32 38  
 FT STRAND 42 47  
 FT TURN 50 51  
 FT STRAND 54 57  
 FT STRAND 63 68  
 FT STRAND 69 74  
 FT HELIX 77 89  
 FT HELIX 90 90  
 FT TURN 95 96  
 FT TURN 99 100  
 FT HELIX 102 104  
 FT TURN 105 108  
 FT TURN 110 111  
 FT HELIX 119 129  
 FT TURN 130 132  
 SQ SEQUENCE 132 AA; 14376 MW; 7CB988AFC2F0B1E4 CRC64;

Query Match 0.8%; Score 7; DB 1; Length 132;  
 Best Local Similarity 100.0%; Pred. No. 31;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 521 APPFOQE 527  
 DB 11 APPFOQE 17

RESULT 13  
 PSBR ARATH STANDARD; PRT; 140 AA.  
 AC P27202;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Photosystem II 10 kDa polypeptide, chloroplast precursor.  
 GN PSBR OR ART1 OR AT1G79040 OR YUP8H12R.29 OR YUP8H12R.34.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosida II; Brassicales; Brassicaceae; Arabidopsids.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=91355945; PubMed=1884004;  
 RA Gil-Gomez G., Marrero P.F., Haro D., Ayte J., Hegardt F.G.;  
 RT "Characterization of the gene encoding the 10 kDa polypeptide of  
 RT photosystem II from Arabidopsis thaliana";  
 RL Plant Mol. Biol. 17:517-522(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RA Raynal M., Grellet F., Laude M., Meyer Y., Cooke R., Delzeny M.;  
 RL Submitted (OCT-1992) to the EMBL/Genbank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=21016719; PubMed=11130712;  
 RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,  
 RA White O., Alonso J., Altsch F., Araujo R., Bowman C.L., Brooks S.Y.,  
 RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,  
 RA Chung M.K., Conn L., Conway A.B., Crosby T.H., Dewar K.,  
 RA Dunn P., Egu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,  
 RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Hultar L.,  
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,  
 RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,  
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,  
 RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maletti R., Marziali A.,  
 RA Miltscher J., Miranda M., Nguyen M., Nieman W.C., Osborne B.I.,  
 RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,  
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,  
 RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,  
 RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,  
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;  
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis  
 RT thaliana";  
 RL Nature 408:816-820(2000).  
 CC -1- FUNCTION: ASSOCIATED WITH THE OXYGEN-EVOLVING COMPLEX OF  
 CC PHOTOSYSTEM II.  
 CC -1- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane; associated  
 CC with the photosystem II complex.  
 CC -1- SIMILARITY: BELONGS TO THE PSBR FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X55970; CAA39441.1; -;  
 DR EMBL: Z17693; CAA79037.1; -;  
 DR EMBL: Z17592; CAA79006.1; -;  
 DR EMBL: AC002986; AAC17052.1; -;  
 DR PIR: S17430; F2MUI0.  
 DR InterPro: IPR006814; Psbr.  
 DR Pfam: PF04725; Psbr.1.  
 KW Photosynthesis; Photosystem II; Chloroplast; Transit peptide;  
 KW Thylakoid; Membrane.  
 FT TRANSIT 1 41 CHLOROPLAST.  
 FT CHAIN 42 140 PHOTOSYSTEM II 10 kDa POLYPEPTIDE.  
 SQ SEQUENCE 140 AA; 14586 MW; ED41DC7E581637EF CRC64;

Query Match 0.8%; Score 7; DB 1; Length 140;  
 Best Local Similarity 100.0%; Pred. No. 33;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 566 RGLPSLT 572  
DB 25 RGLPSLT 31

RESULT 14  
PSBR BRACM STANDARD; PRT; 141 AA.

AC P49108;  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Photosystem II 10 kDa polypeptide, chloroplast precursor.  
GN PSBR.

OS Brassica campestris (Field mustard).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC euroside II; Brassicales; Brassicaceae; Brassica.  
OX NCBI\_TaxId=3711;  
[1]  
RN SEQUENCE FROM N.A.  
RC STRAIN=cv. Pekinensis;  
RX MEDLINE=95334521; PubMed=7610192;  
RA Kim H.U., Yun C.H., Park B.S., Ryu J.C., Chung T.Y.;  
RT "Nucleotide sequence of a cDNA clone encoding the complete precursor  
for the '10-kilodalton' polypeptide of photosystem II from Chinese  
cabbage.";  
RL Plant Physiol. 108:867-867(1995).

CC -1- FUNCTION: ASSOCIATED WITH THE OXYGEN-EVOLVING COMPLEX OF  
PHOTOSYSTEM II.  
CC -1- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane; associated  
with the photosystem II complex.  
CC -1- SIMILARITY: BELONGS TO THE PSBR FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: L31936; AAA74957.1; -  
DR InterPro: IPR006814; Psbr.  
DR Pfam: PF04725; Psbr; 1.  
KW Photosynthesis; Photosystem II; Chloroplast; Transit peptide;  
KW Thylakoid; Membrane.  
FT TRANSIT 1 42 CHLOROPLAST (BY SIMILARITY).  
FT CHAIN 43 141 PHOTOSYSTEM II 10 KDA POLYPEPTIDE.  
SQ SEQUENCE 141 AA; 14649 MW; AA85D2561ECFEFA CRC64;

Query Match 0.8%; Score 7; DB 1; Length 141;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 566 RGLPSLT 572  
DB 25 RGLPSLT 31

RESULT 15  
MOAC METUA STANDARD; PRT; 152 AA.

AC Q58535;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Probable molybdenum cofactor biosynthesis protein C.  
GN MOAC OR MJ1135.  
OS Methanococcus jannaschii.  
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;  
OC Methanocaldococcales; Methanocaldococcus.

OX NCBI\_TaxId=2190;  
RN (1)  
RP SEQUENCE FROM N.A.  
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;  
RX MEDLINE=96337999; PubMed=8668087;  
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,  
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,  
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,  
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,  
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhmann J.L., Nguyen D.,  
RA Uiterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,  
RA Catterton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,  
RA Kleink H.-P., Frazer C.M., Smith H.O., Woese C.R., Venter J.C.;  
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus  
jannaschii.";  
RL Science 273:1058-1073(1996).

CC -1- FUNCTION: Together with moaA, is involved in the conversion of a  
guanosine derivative (GXP) into molybdopterin precursor Z (By  
similarity).  
CC -1- PATHWAY: Molybdenum cofactor biosynthesis; first step.  
CC -1- SIMILARITY: Belongs to the moaC family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: U67556; AAB99137.1; -  
DR PIR: F64441; F64441.  
DR HSSP: P30747; 1EKR.  
DR TIGR: MJ1135; -  
DR HAMAP: MF\_01224; -; 1.  
DR InterPro: IPR002820; Moac.  
DR Pfam: PF01967; Moac; 1.  
DR TIGRFAMs: TIGR00581; moac; 1.  
KW Molybdenum cofactor biosynthesis; Complete proteome.  
FT ACT SITE 123 POTENTIAL.  
SQ SEQUENCE 152 AA; 17043 MW; D91AC8BD1D2852AF CRC64;

Query Match 0.8%; Score 7; DB 1; Length 152;  
Best Local Similarity 100.0%; Pred. No. 35;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GVSIALL 13  
DB 113 GVSIALL 119

Search completed: February 20, 2004, 19:47:10  
Job time : 29 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: February 20, 2004, 18:39:32 ; Search time 119 Seconds

(without alignments)  
1812.874 Million cell updates/sec

Title: US-09-899-569a-4

Perfect score: 836  
Sequence: 1 MAGNCGVSIALGVLLGA.....SSKTDIPLNTQEPMPAE 836

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 830525 seqs, 258052604 residues

Word size : 0

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

SPTREMBL\_23:\*

1: sp\_archaea:\*

2: sp\_bacteria:\*

3: sp\_fungi:\*

4: sp\_human:\*

5: sp\_invertebrate:\*

6: sp\_mammal:\*

7: sp\_mmc:\*

8: sp\_organelle:\*

9: sp\_phage:\*

10: sp\_plant:\*

11: sp\_rodent:\*

12: sp\_virus:\*

13: sp\_vertebrate:\*

14: sp\_unclassified:\*

15: sp\_virus:\*

16: sp\_bacterioph:\*

17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	836	100.0	836	4	Q96GQ7
2	725	86.7	836	4	Q9H5V8
3	382	45.7	649	4	Q9H8C2
4	341	40.8	343	4	Q8WU91
5	291	34.8	392	4	Q9H676
6	39	4.7	320	11	Q8K246
7	39	4.7	470	11	Q8BZD9
8	27	3.2	384	11	Q921W9
9	9	1.1	108	5	Q96238
10	11	1.1	815	5	Q23860
11	9	1.1	1245	5	Q812R3
12	9	1.1	1629	5	Q9U0K9
13	9	1.1	2747	5	Q81C15
14	8	1.0	108	11	Q8BR25
15	8	1.0	127	11	Q8C3C4
16	8	1.0	144	10	Q9LEZ1

17	8	1.0	173	8	Q8HKR3	Q8HKR3 arcos sp. k
18	8	1.0	179	4	Q9UHQ2	Q9UHQ2 homo sapien
19	8	1.0	187	4	Q8X6M2	Q8X6M2 escherichia
20	8	1.0	190	4	Q96S03	Q96S03 homo sapien
21	8	1.0	195	16	Q8CVR9	Q8CVR9 escherichia
22	8	1.0	203	5	Q17789	Q17789 caenorhabdi
23	8	1.0	205	17	Q8TVN0	Q8TVN0 methanopyru
24	8	1.0	206	16	Q8DM11	Q8DM11 synechococc
25	8	1.0	221	4	Q9BRV3	Q9BRV3 homo sapien
26	8	1.0	221	4	Q9UHQ3	Q9UHQ3 homo sapien
27	8	1.0	221	6	Q9SKW8	Q9SKW8 papio anubi
28	8	1.0	232	16	Q8ZS89	Q8ZS89 anabasa sp
29	8	1.0	262	16	Q9KD01	Q9KD01 bacillus ha
30	8	1.0	285	5	Q9U6J3	Q9U6J3 trichomegal
31	8	1.0	349	5	Q9BIJ3	Q9BIJ3 sarcocystis
32	8	1.0	380	10	Q82468	Q82468 mesembryant
33	8	1.0	390	3	Q9USL5	Q9USL5 schizosach
34	8	1.0	443	2	Q8RTV3	Q8RTV3 uncultured
35	8	1.0	468	3	Q9P3W6	Q9P3W6 schizosach
36	8	1.0	484	5	Q9VAF4	Q9VAF4 drosophila
37	8	1.0	484	11	Q8BJ24	Q8BJ24 mus musculu
38	8	1.0	509	5	Q9B1D1	Q9B1D1 anophelis g
39	8	1.0	510	10	Q8W2M4	Q8W2M4 nicotiana t
40	8	1.0	550	2	Q9RN37	Q9RN37 salmoneila
41	8	1.0	565	5	Q81151	Q81151 plasmodium
42	8	1.0	597	11	Q9JKC1	Q9JKC1 ratius norv
43	8	1.0	682	16	Q9HYO5	Q9HYO5 pseudomonas
44	8	1.0	899	5	Q97296	Q97296 plasmodium
45	8	1.0	1137	11	Q8K561	Q8K561 mus musculu

## ALIGNMENTS

### RESULT 1

ID Q96GQ7 PRELIMINARY; PRT; 836 AA.

AC Q96GQ7;

DT 01-DEC-2001 (TREMBLrel. 19, Created)

DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE CUB domain containing protein 1.

GN CDCP1.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OC NCBI\_TaxID=9606;

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=21359860; PubMed=11466621;

RA Scherl-Mostaguer M., Sommergruber W., Abseher R., Hauptmann R., Ambros P., Schweifer N.;

RT "Identification of a novel gene, CDCP1, overexpressed in human colorectal cancer."

RT Oncogene 20:4402-4408(2001).

RL EMBL: AY026461; AAK02058.1;

DR EMBL: 836 AA; 92873 MW; FB4D2DBDD35C519 CRC64;

SQ SEQUENCE

QY	Query Match	100.0%;	Score 836;	DB 4;	Length 836;
QY	Best Local Similarity	100.0%;	Pred. No. 0;		
QY	Matches 836;	Conservative	0;	Mismatches	0;
QY				Indels	0;
QY				Gaps	0;
QY	1	MAGNCGVSIALGVLLGAARLP	PGAAAFETALPRESNITVLTGPTTLAKCYVI	60	
QY	1	MAGNCGVSIALGVLLGAARLP	PGAAAFETALPRESNITVLTGPTTLAKCYVI	60	
QY	61	SKRHITMISIGERIVTFSCQSPENH	FVIEIQNIDCMGSPCFGVOLOPSTSLPT	120	
QY	61	SKRHITMISIGERIVTFSCQSPENH	FVIEIQNIDCMGSPCFGVOLOPSTSLPT	120	
QY	121	LNRFTIWDVKAHKSIGLELOP	SIPRLQIGPESCPDGVTHSISGRIDATVVRIGTFCSN	180	
QY	121	LNRFTIWDVKAHKSIGLELOP	SIPRLQIGPESCPDGVTHSISGRIDATVVRIGTFCSN	180	
QY	121	LNRFTIWDVKAHKSIGLELOP	SIPRLQIGPESCPDGVTHSISGRIDATVVRIGTFCSN	180	

```
QY 181 GTVSRIKMOEGVYKMAHLHPMFHPRNVSGFSIANRSSIKRLCTIESEVEGEGSATLMSANY 240
DB 181 GTVSRIKMOEGVYKMAHLHPMFHPRNVSGFSIANRSSIKRLCTIESEVEGEGSATLMSANY 240
QY 241 PGCFPEDELMTMOQFVPAHLRASVSFLNPNLSNCRKEERVEYY1PGSTTNEVPFLBCK 300
DB 241 PGCFPEDELMTMOQFVPAHLRASVSFLNPNLSNCRKEERVEYY1PGSTTNEVPFLBCK 300
QY 301 QPGNAGNFNLISLQGCDDAOSPGILRLQFVLYVHPQNESNKIYVVDLSNERAMSLTIE 360
DB 301 QPGNAGNFNLISLQGCDDAOSPGILRLQFVLYVHPQNESNKIYVVDLSNERAMSLTIE 360
QY 361 PRPVQSRKFVPGCFVCLSESRITSSNLTLTSGSKHISFLCDLTRLMMNVKTTISCTDH 420
DB 361 PRPVQSRKFVPGCFVCLSESRITSSNLTLTSGSKHISFLCDLTRLMMNVKTTISCTDH 420
QY 421 RYCQRKSYSLQVPSDILHPVELHDFSMKLVKPKRLSLVPAQKLOQTHEKPCNTSF 480
DB 421 RYCQRKSYSLQVPSDILHPVELHDFSMKLVKPKRLSLVPAQKLOQTHEKPCNTSF 480
QY 481 SYLVASAIPSODLYGSGFCPGSGIKQIOVKONISVTLRTFAPSFOEASRQGLTVSFIFY 540
DB 481 SYLVASAIPSODLYGSGFCPGSGIKQIOVKONISVTLRTFAPSFOEASRQGLTVSFIFY 540
QY 541 FKEBGFVTVPTDKSKVYLRTIPNMDRGLPSLTSVSNISVPRDQVACLTFEKRSGVVCQ 600
DB 541 FKEBGFVTVPTDKSKVYLRTIPNMDRGLPSLTSVSNISVPRDQVACLTFEKRSGVVCQ 600
QY 601 TGRAPMIIQEOQTRAEIISLDEVDLPKPSFHHHSFWNINISCSPTSGQDLPLFSVTLT 660
DB 601 TGRAPMIIQEOQTRAEIISLDEVDLPKPSFHHHSFWNINISCSPTSGQDLPLFSVTLT 660
QY 661 PRVVDLTVILIAAVGGVLLLSALGLIICVYKSKKKTKNKGPVAGIYNGNINTEMPROK 720
DB 661 PRVVDLTVILIAAVGGVLLLSALGLIICVYKSKKKTKNKGPVAGIYNGNINTEMPROK 720
QY 721 KFQGRKNDSHVAVIEDTMVYGHLLQDSSGSLQPEVDIYRPPQGTMGVCPSPPTIC 780
DB 721 KFQGRKNDSHVAVIEDTMVYGHLLQDSSGSLQPEVDIYRPPQGTMGVCPSPPTIC 780
QY 781 SRAPFAKLTATEPPRSPSESESEPTFSHPNNGVSSKDTDIPLNTQEPMEPAE 836
DB 781 SRAPFAKLTATEPPRSPSESESEPTFSHPNNGVSSKDTDIPLNTQEPMEPAE 836
```

## RESULT 2

```
Q9H5V8 PRELIMINARY; PRT; 836 AA.
ID Q9H5V8;
AC Q9H5V8;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein FLJ22969.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCB1_TaxID=9606;
RN [1]_
RP SEQUENCE FROM N.A.
RA Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,
RA Suzuki Y., Ohtsuka M., Nishi T., Shibahara T., Tanaka T.,
RA Nakamura Y., Iwagaki T., Sugano S.,
RT "NBOO human cDNA sequencing project."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK026622; BAB1511.1; -.
KW Hypothetical protein.
```

Query Match 86.7%; Score 725; DB 4; Length 836;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 825; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 MAGLNGCVSIALGLVLLGAARLPRGAFAFIALPRESNTIYLIKGLPTLLAKCYIYI 60
DB 1 MAGLNGCVSIALGLVLLGAARLPRGAFAFIALPRESNTIYLIKGLPTLLAKCYIYI 60
QY 61 SKRHITMLISGSEIVTFPSCQSPENHFVIEIQKIDCMSCPCPFGEVQLOPSTSLPT 120
DB 61 SKRHITMLISGSEIVTFPSCQSPENHFVIEIQKIDCMSCPCPFGEVQLOPSTSLPT 120
QY 121 LNRTFIWVKAKHSIGLELOPSIPLRLQIGPESCPDGVTHISGRIDATVVRIGTFCGN 180
DB 121 LNRTFIWVKAKHSIGLELOPSIPLRLQIGPESCPDGVTHISGRIDATVVRIGTFCGN 180
QY 181 GTVSRIKMOEGVYKMAHLHPMFHPRNVSGFSIANRSSIKRLCTIESEVEGEGSATLMSANY 240
DB 181 GTVSRIKMOEGVYKMAHLHPMFHPRNVSGFSIANRSSIKRLCTIESEVEGEGSATLMSANY 240
QY 241 PGCFPEDELMTMOQFVPAHLRASVSFLNPNLSNCRKEERVEYY1PGSTTNEVPFLBCK 300
DB 241 PGCFPEDELMTMOQFVPAHLRASVSFLNPNLSNCRKEERVEYY1PGSTTNEVPFLBCK 300
QY 301 QPGNAGNFNLISLQGCDDAOSPGILRLQFVLYVHPQNESNKIYVVDLSNERAMSLTIE 360
DB 301 QPGNAGNFNLISLQGCDDAOSPGILRLQFVLYVHPQNESNKIYVVDLSNERAMSLTIE 360
QY 361 PRPVQSRKFVPGCFVCLSESRITSSNLTLTSGSKHISFLCDLTRLMMNVKTTISCTDH 420
DB 361 PRPVQSRKFVPGCFVCLSESRITSSNLTLTSGSKHISFLCDLTRLMMNVKTTISCTDH 420
QY 421 RYCQRKSYSLQVPSDILHPVELHDFSMKLVKPKRLSLVPAQKLOQTHEKPCNTSF 480
DB 421 RYCQRKSYSLQVPSDILHPVELHDFSMKLVKPKRLSLVPAQKLOQTHEKPCNTSF 480
QY 481 SYLVASAIPSODLYGSGFCPGSGIKQIOVKONISVTLRTFAPSFOEASRQGLTVSFIFY 540
DB 481 SYLVASAIPSODLYGSGFCPGSGIKQIOVKONISVTLRTFAPSFOEASRQGLTVSFIFY 540
QY 541 FKEBGFVTVPTDKSKVYLRTIPNMDRGLPSLTSVSNISVPRDQVACLTFEKRSGVVCQ 600
DB 541 FKEBGFVTVPTDKSKVYLRTIPNMDRGLPSLTSVSNISVPRDQVACLTFEKRSGVVCQ 600
QY 601 TGRAPMIIQEOQTRAEIISLDEVDLPKPSFHHHSFWNINISCSPTSGQDLPLFSVTLT 660
DB 601 TGRAPMIIQEOQTRAEIISLDEVDLPKPSFHHHSFWNINISCSPTSGQDLPLFSVTLT 660
QY 661 PRVVDLTVILIAAVGGVLLLSALGLIICVYKSKKKTKNKGPVAGIYNGNINTEMPROK 720
DB 661 PRVVDLTVILIAAVGGVLLLSALGLIICVYKSKKKTKNKGPVAGIYNGNINTEMPROK 720
QY 721 KFQGRKNDSHVAVIEDTMVYGHLLQDSSGSLQPEVDIYRPPQGTMGVCPSPPTIC 780
DB 721 KFQGRKNDSHVAVIEDTMVYGHLLQDSSGSLQPEVDIYRPPQGTMGVCPSPPTIC 780
QY 781 SRAPFAKLTATEPPRSPSESESEPTFSHPNNGVSSKDTDIPLNTQEPMEPAE 836
DB 781 SRAPFAKLTATEPPRSPSESESEPTFSHPNNGVSSKDTDIPLNTQEPMEPAE 836
```

## RESULT 3

```
Q9H8C2 PRELIMINARY; PRT; 649 AA.
ID Q9H8C2;
AC Q9H8C2;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein FLJ13772.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCB1_TaxID=9606;
RN [1]_
RP SEQUENCE FROM N.A.
RA Tissue=Placenta;
RA Iwagaki T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
```

RA Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,  
RA Tanase T., Nomura Y., Togiyasu S., Komai F., Hara R., Takeuchi K.,  
RA Arima M., Nabebara T., Ishii S., Kawai Y., Saito K., Yamamoto J.,  
RA Makamatsu A., Nakamura Y., Nagahara K., Masuo Y., Oshima A.,  
RT "NEDO human cDNA sequencing project."  
RL Submitted (AUG-2000) to the EMBL/Genbank/DBJ databases.  
DR EMBL; AK023834; BAB14695.1; -  
KW Hypothetical protein.  
SQ SEQUENCE 649 AA; 72579 MW; 0F4404AD31D60E9A CRC64;

Query Match 45.7%; Score 382; DB 4; Length 649;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 582; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 253 QFVPAHLRAVSFLNPLNSNCERKEVEYIIPGTTNPPEVKLEDKQPMAGNPLS 312  
DB 66 QFVPAHLRAVSFLNPLNSNCERKEVEYIIPGTTNPPEVKLEDKQPMAGNPLS 125  
QY 313 LOGCDQDQSPGILRLQFOVLVQHPQONESNKIYVVDLSNERAMSITTEPRVKSKEFP 372  
DB 126 LOGCDQDQSPGILRLQFOVLVQHPQONESNKIYVVDLSNERAMSITTEPRVKSKEFP 185  
QY 373 GCFVCLSRSTSSNLTLSGSKHISFLCDLTRLMMNVETISCTDHRVYCKRSYSLOV 432  
DB 186 GCFVCLSRSTSSNLTLSGSKHISFLCDLTRLMMNVETISCTDHRVYCKRSYSLOV 245  
QY 433 PSDILHLPEVLHDFSWKLVLPKDRSLVLPAPKQLQHTHEKPCNTSPSYLVASIPSD 492  
DB 246 PGDILHLPEVLHDFSWKLVLPKDRSLVLPAPKQLQHTHEKPCNTSPSYLVASIPSD 305.  
QY 493 LYFGSFCGSGSIKIQVKNISVTLRTAPSPFOEASRQGLTVSFIPIYKKEGVFTVLPD 552  
DB 306 LYFGSFCGSGSIKIQVKNISVTLRTAPSPFOEASRQGLTVSFIPIYKKEGVFTVLPD 365  
QY 553 TKSRYLRTPMWDGSLPGLSVSNISVPRDQVACLTFKRSQVCTGAPMIOEOR 612  
DB 366 TKSRYLRTPMWDGSLPGLSVSNISVPRDQVACLTFKRSQVCTGAPMIOEOR 425  
QY 613 TRAEIFSLDEVDLPKPSFHHSFVWVINSNCSPTSGKDLILFSVTLRPRVLDLVILIA 672  
DB 426 TRAEIFSLDEVDLPKPSFHHSFVWVINSNCSPTSGKDLILFSVTLRPRVLDLVILIA 485  
QY 673 AVGGVLLLSLGLIICVKKKKKKTKNGPAVINGNINTEMPROPKPKQGRKNDSH 732  
DB 486 AVGGVLLLSLGLIICVKKKKKKTKNGPAVINGNINTEMPROPKPKQGRKNDSH 545  
QY 733 VYAVIEDTMYGHLIQQSSGSLFQPEVDTPRPFGTGWCVCPSPPTICSRAPLATEE 792  
DB 546 VYAVIEDTMYGHLIQQSSGSLFQPEVDTPRPFGTGWCVCPSPPTICSRAPLATEE 605  
QY 793 PPRSPPESESEPTESHNNGDVSSKOTDIPLANTOPEMPEAE 836  
DB 606 PPRSPPESESEPTESHNNGDVSSKOTDIPLANTOPEMPEAE 649

RESULT 4

Q8WU91 PRELIMINARY; PRT; 343 AA.  
AC Q8WU91;  
DT 01-MAR-2002 (TREMblrel. 20, Created)  
DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)  
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)  
DE Similar to hypothetical protein FLJ22969.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Kidney;  
RA Strauberg R.;  
RL Submitted (JAN-2002) to the EMBL/Genbank/DBJ databases.  
DR EMBL; BC021099; AAH21099.1; -

KW Hypothetical protein.  
SQ SEQUENCE 343 AA; 37817 MW; 3E4E13379D9D94D1B CRC64;

Query Match 40.8%; Score 341; DB 4; Length 343;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGNLCVSIALLVLLIGARLPRGAFAFIALPRESNTIVLIGTPTLLAKPCYVI 60  
DB 1 MAGNLCVSIALLVLLIGARLPRGAFAFIALPRESNTIVLIGTPTLLAKPCYVI 60  
QY 61 SKRIITMLSIKSGRIYFTFSCQSPENHVEIIOKNIDCMGCPGPGVQLOPSTSLPT 120  
DB 61 SKRIITMLSIKSGRIYFTFSCQSPENHVEIIOKNIDCMGCPGPGVQLOPSTSLPT 120  
QY 121 LNRFTIWDVKAHKSIGLELOFSIRLNOIGBSCPCGVTHSISGRIDATVVRIGTFCSN 180  
DB 121 LNRFTIWDVKAHKSIGLELOFSIRLNOIGBSCPCGVTHSISGRIDATVVRIGTFCSN 180  
QY 181 GTVARIKMQEGVKAALHLPFHPNVSQFSIANSISIKRLCTISVFEGBSATLMSANY 240  
DB 181 GTVARIKMQEGVKAALHLPFHPNVSQFSIANSISIKRLCTISVFEGBSATLMSANY 240  
QY 241 PEGPPEDELMTWQFVPAHLRAVSFLNPLNSNCERKEVEYIIPGTTNPPEVKLEDK 300  
DB 241 PEGPPEDELMTWQFVPAHLRAVSFLNPLNSNCERKEVEYIIPGTTNPPEVKLEDK 300  
QY 301 QPGNMGAFNLSLOGCDQDQSPGILRLQFOVLVQHPONES 341  
DB 301 QPGNMGAFNLSLOGCDQDQSPGILRLQFOVLVQHPONES 341

RESULT 5

Q9H676 PRELIMINARY; PRT; 392 AA.  
AC Q9H676;  
DT 01-MAR-2001 (TREMblrel. 16, Created)  
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)  
DT 01-OCT-2002 (TREMblrel. 22, Last annotation update)  
DE Hypothetical protein FLJ22534 (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,  
RA Oktani R., Ota T., Suzuki Y., Odayashi M., Nishi T., Shibahara T.,  
RA Tanaka T., Nakamura Y., Isegaki T., Sugano S.,  
RT "NEDO human cDNA sequencing project."  
RL Submitted (AUG-2000) to the EMBL/Genbank/DBJ databases.  
DR EMBL; AK026187; BAB15388.1; -  
KW Hypothetical protein.  
FT NON TER 392  
SQ SEQUENCE 392 AA; 44031 MW; 3E21C8B2703E52F3 CRC64;

Query Match 34.8%; Score 291; DB 4; Length 392;  
Best Local Similarity 99.7%; Pred. No. 6,4e-290;  
Matches 391; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 305 MAGNPNLSLOGCDQDQSPGILRLQFOVLVQHPONESNKIYVVDLSNERAMSITTEPRV 364  
DB 1 MAGNPNLSLOGCDQDQSPGILRLQFOVLVQHPONESNKIYVVDLSNERAMSITTEPRV 60  
QY 365 KQSRKFPVGCFCVCLSRSTSSNLTLSGSKHISFLCDLTRLMMNVETISCTDHRVYCO 424  
DB 61 KQSRKFPVGCFCVCLSRSTSSNLTLSGSKHISFLCDLTRLMMNVETISCTDHRVYCO 120  
QY 425 RKYSYLOVPSDILHLPEVLHDFSWKLVLPKDRSLVLPAPKQLQHTHEKPCNTSPSYLV 484  
DB 121 RKYSYLOVPSDILHLPEVLHDFSWKLVLPKDRSLVLPAPKQLQHTHEKPCNTSPSYLV 180  
QY 485 ASAIPSDLYFGSFCGSGSIKIQVKNISVTLRTAPSPFOEASRQGLTVSFIPIYKKE 544

DB 181 AAIASQDILYFSGFCGGSIKQIQVKNQISVTLRTFAPSFQOASRQGLTVSFIPFKKE 240  
QY 545 GVFVTVPDTRSKYVLTPTMMDRGLPBLTISVSNISVPRQVACTFFPKERSGVWCOTGRA 604  
DB 241 GVFVTVPDTRSKYVLTPTMMDRGLPBLTISVSNISVPRQVACTFFPKERSGVWCOTGRA 300  
QY 605 FMIIOEORTAEIISLDEDLVLPKPSFHHHSFWNVINSNCSPTSGQDLILFSVTLTPRTV 664  
DB 301 FMIIOEORTAEIISLDEDLVLPKPSFHHHSFWNVINSNCSPTSGQDLILFSVTLTPRTV 360  
QY 665 DLTVLIAVGGVLLSLAGLITICVKKKK 696  
DB 361 DLTVLIAVGGVLLSLAGLITICVKKKK 392

RESULT 6  
QY 08K246 PRELIMINARY; PRT; 320 AA.  
AC 08K246;  
DT 01-OCT-2002 (TREMBLrel. 22, Created)  
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)  
DE 01-OCT-2002 (TREMBLrel. 22, Last annotation update)  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Strausberg R.;  
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC034137; AAH34137.1; -  
KW Hypothetical protein.  
SQ SEQUENCE 320 AA; 35962 MW; 050906547833BCG6 CRC64;

Query Match 4.7%; Score 39; DB 11; Length 320;  
Best Local Similarity 100.0%; Pred. No. 5.8e-31;  
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 297 LEDKOPGNMAGNPNLSLOGCDODASPGILRLQFOVLVQ 335  
DB 110 LEDKOPGNMAGNPNLSLOGCDODASPGILRLQFOVLVQ 148

RESULT 7  
QY 08B2D9 PRELIMINARY; PRT; 470 AA.  
AC 08B2D9;  
DT 01-MAR-2003 (TREMBLrel. 23, Created)  
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE Similar to CUB domain containing protein 1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA STRAIN=C57BL/6J; TISSUE=urinary bladder;  
RX MEDLINE=22354683; PubMed=12466851;  
RA The FANTOM Consortium;  
RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs.";  
RL Nature 420:563-573 (2002).  
DR EMBL; AUC035674; BAC29147.1; -  
SQ SEQUENCE 470 AA; 52277 MW; A36CD2B547348383 CRC64;

Query Match 4.7%; Score 39; DB 11; Length 470;  
Best Local Similarity 100.0%; Pred. No. 8.1e-31;  
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 297 LEDKOPGNMAGNPNLSLOGCDODASPGILRLQFOVLVQ 335  
DB 297 LEDKOPGNMAGNPNLSLOGCDODASPGILRLQFOVLVQ 335

RESULT 8  
QY 0921M9 PRELIMINARY; PRT; 384 AA.  
AC 0921M9;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE Unknown (Protein for IMAGE:3709937) (Fragment).  
GN 9030022E12R1K OR AA409659.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Strausberg R.;  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC011340; AAH11340.1; -  
DR MGD; MGI:1925928; 9030022E12R1K.  
FT NON\_TER 1  
SQ SEQUENCE 384 AA; 42406 MW; 4B5F17D30FE16080 CRC64;

Query Match 3.2%; Score 27; DB 11; Length 384;  
Best Local Similarity 100.0%; Pred. No. 1.5e-18;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 724 KGRKNDSHVAVIEDTWYGHLLQDS 750  
DB 272 KGRKNDSHVAVIEDTWYGHLLQDS 298

RESULT 9  
QY 096238 PRELIMINARY; PRT; 108 AA.  
AC 096238;  
DT 01-MAY-1999 (TREMBLrel. 10, Created)  
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE Hypothetical protein.  
GN PFB0725C.  
OS Plasmodium falciparum (Isolate 3D7).  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI\_TaxID=36329;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA STRAIN=3D7;  
RX MEDLINE=99021743; PubMed=9804551;  
RA Gardner M.J.; Tetstelin H.; Carucci D.J.; Cummings L.M.; Aravind L.;  
RA Koonin E.V.; Shallow S.; Mason T.; Yu K.; Fujii C.; Pederson J.;  
RA Shen K.; Jing J.; Ascon C.; Lai Z.; Schwartz D.C.; Petrea M.;  
RA Salberg S.; Zhou L.; Sutton G.G.; Clayton R.; White H.O.;  
RA Fraser C.M.; Adams M.D.; Venter J.C.; Hoffman S.L.;  
RT "Chromosome 2 sequence of the human malaria parasite Plasmodium  
RT falciparum.";  
RL Science 282:1126-1132 (1998).  
RN [2]  
RP SEQUENCE FROM N.A.

RA STRAIN=3D7;  
RX MEDLINE=22255705; PubMed=12368864;  
RA Gardner M.J.; Hall N.; Fung E.; White O.; Berriman M.; Hyman R.W.;  
RA Carlton J.M.; Pain A.; Nelson K.E.; Bowman S.; Paulsen I.T.; James K.;  
RA Eisen J.A.; Nene V.; Rutherford K.; Salzberg S.L.; Craig A.; Kyse S.;  
RA Chan M.-S.; Nene V.; Shallow S.J.; Sub B.; Peterson J.; Angiolini S.;  
RA Pertea M.; Allen J.; Selengut J.; Haft D.; Mather M.W.; Vaidya A.B.;  
RA Martin D.M.A.; Fairlamb A.H.; Fraunholz M.J.; Roos D.S.; Ralph S.A.;  
RA McFadden G.I.; Cummings L.M.; Subramanian G.M.; Mungall C.;  
RA Venter J.C.; Carucci D.J.; Hoffman S.L.; Newbold C.; Davis R.W.;  
RA Fraser C.M.; Bartell B.;

RT "Genome sequence of the human malaria parasite Plasmodium  
 RT falciparum.";  
 RL Nature 419:498-511(2002).  
 DR EMBL; AB01414; AAC71934.2; -  
 KW Hypothetical protein.  
 SQ SEQUENCE 108 AA; 13675 MW; 11DC06350C0362B5 CRC64;  
 Query Match 1.1%; Score 9; DB 5; Length 108;  
 Best Local Similarity 100.0%; Pred. No. 1.5;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 692 KKKKKKTKN 700  
 DB 100 KKKKKKTKN 108  
 RESULT 10  
 ID Q23860 PRELIMINARY; PRT; 815 AA.  
 AC Q23860;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE Vacuolar proton ATPase 100-kDa subunit.  
 GN VATM.  
 OS Dictyostelium discoideum (Slime mold).  
 CC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.  
 CX NCBI\_TaxID=44689;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96338613; PubMed=8743951;  
 RA Liu T., Clarke M.;  
 DR EMBL; U38803; AAB49621.1; -  
 DR InterPro; IPR002490; V\_ATPase\_sub116.  
 DR Pfam; PF01496; V\_ATPase\_sub\_a; 1.  
 SQ SEQUENCE 815 AA; 93272 MW; DF0858C79F9F75D2 CRC64;  
 Query Match 1.1%; Score 9; DB 5; Length 815;  
 Best Local Similarity 100.0%; Pred. No. 9.1;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 679 LLLSALGLI 687  
 DB 430 LLLSALGLI 438  
 RESULT 11  
 ID O812R3 PRELIMINARY; PRT; 1245 AA.  
 AC O812R3;  
 DT 01-MAR-2003 (TREMBlrel. 23, Created)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE Hypothetical protein.  
 GN PF1120W.  
 OS Plasmodium falciparum (isolate 3D7).  
 CC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 CX NCBI\_TaxID=36329;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22255708; PubMed=13368867;  
 RA Hall N., Pain A., Berriman M., Churcher C., Harris B., Harris D.,  
 RA Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brooks K.,  
 RA Buceae C.O., Burrows C., Cherevach I., Chillingworth C.,  
 RA Chillingworth T., Christodoulou Z., Clark L., Clark R., Corton C.,  
 RA Cronin A., Davies R., Davis P., Dear P., Dearden F., Doggett J.,  
 RA Feltwell T., Goble A., Goodhead I., Gilliam R., Hamlin N., Hance Z.,  
 RA Harper D., Hauser H., Hornsby T., Holtroyd S., Horrocks P.,  
 RA Humphrey S., Jagsis K., James K.D., Johnson D., Kerhornou A.,  
 RA Knights A., Konfortov B., Kyes S., Larke N., Lawson D., Lennard N.,

RA Line A., Maddison M., McLean J., Mooney P., Moule S., Murphy L.,  
 RA Oliver K., Ormond D., Price C., Quail M.A., Rabinowitsch E.,  
 RA Rajendram M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M.,  
 RA Seeger K., Sharp S., Smith R., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J.,  
 RA Sultson J.E., Craig A., Newbold C., Barrell B.G.;  
 RT "Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13."  
 RL Nature 419:527-531(2002).  
 DR EMBL; AL929357; CAD51926.1; -  
 KW Hypothetical protein.  
 SQ SEQUENCE 1245 AA; 144680 MW; 12904C239AED17F1 CRC64;  
 Query Match 1.1%; Score 9; DB 5; Length 1245;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 692 KKKKKKTKN 700  
 DB 690 KKKKKKTKN 698  
 RESULT 12  
 ID Q9UOK9 PRELIMINARY; PRT; 1629 AA.  
 AC Q9UOK9;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)  
 DE Hypothetical 197.4 kDa protein.  
 GN MALP2.13.  
 OS Plasmodium falciparum (isolate 3D7).  
 CC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 CX NCBI\_TaxID=36329;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=3D7;  
 RA Devlin K., Bowman S., Churcher C., Harris B., Harris D., Lawson D.,  
 RA Quail M., Barrell B.;  
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL035475; CAB62854.2; -  
 KW Hypothetical protein.  
 SQ SEQUENCE 1629 AA; 197412 MW; C1233E537648C9D CRC64;  
 Query Match 1.1%; Score 9; DB 5; Length 1629;  
 Best Local Similarity 100.0%; Pred. No. 17;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 692 KKKKKKTKN 700  
 DB 498 KKKKKKTKN 506  
 RESULT 13  
 ID O81CL5 PRELIMINARY; PRT; 2747 AA.  
 AC O81CL5;  
 DT 01-MAR-2003 (TREMBlrel. 23, Created)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE Ribonuclease, putative.  
 GN MAL6P1.272.  
 OS Plasmodium falciparum (isolate 3D7).  
 CC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 CX NCBI\_TaxID=36329;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Cherevach I., Davis P., Goodhead I., Stevens K., Mungall K.,  
 RA Berriman M., Pain A., Hall N., Atkin R., Chillingworth C., Doggett J.,  
 RA Ormond D., Sanders M., Hayes R., Hall S., Quail M., Barrell B.,  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL844505; CAD50426.1; -  
 SQ SEQUENCE 2747 AA; 325595 MW; F01B08D441D0D3E5 CRC64;

Query Match 1.1%; Score 9; DB 5; Length 2747;  
 Best Local Similarity 100.0%; Pred. No. 27;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 29 KKKKTKNK 36  
 Search completed: February 20, 2004, 19:49:22  
 Job time : 123 secs

RESULT 14

Q8BR25 PRELIMINARY; PRT; 108 AA.  
 AC Q8BR25;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Hypothetical protein.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Brain;  
 RX MEDLINE=22354683; PubMed=12466851;  
 RA The FANTOM Consortium.  
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs.";  
 RL Nature 420:563-573(2002).  
 DR EMBL; AK045823; BAC32505.1; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 108 AA; 12201 MW; A9F6B7CC972EA55B CRC64;

Query Match 1.0%; Score 8; DB 11; Length 108;  
 Best Local Similarity 100.0%; Pred. No. 16;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 693 KKKKTKNK 700  
 DB 10 KKKKTKNK 17

RESULT 15

Q8C3C4 PRELIMINARY; PRT; 127 AA.  
 ID Q8C3C4  
 AC Q8C3C4;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Hypothetical protein.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Head;  
 RX MEDLINE=22354683; PubMed=12466851;  
 RA The FANTOM Consortium.  
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs.";  
 RL Nature 420:563-573(2002).  
 DR EMBL; AK086308; BAC39646.1; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 127 AA; 14090 MW; 24E2FC1C01F1EA2F CRC64;

Query Match 1.0%; Score 8; DB 11; Length 127;  
 Best Local Similarity 100.0%; Pred. No. 19;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 693 KKKKTKNK 700



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2004, 19:45:04 ; Search time 32 Seconds  
(without alignments)  
1105.371 Million cell updates/sec

Title: US-09-899-569A-4  
Perfect score: 836  
Sequence: 1 MAGNCGVSIALGVLLGA.....SSKTDIPLNTQEPMEPAE 836

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 328717 seqs, 42310858 residues

Word size : 0

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCFUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	413	49.4	414	US-09-489-847-273	Sequence 273, App
2	413	49.4	443	US-09-489-847-271	Sequence 271, App
3	211	25.2	709	US-09-489-847-132	Sequence 132, App
4	159	19.0	159	US-09-489-847-272	Sequence 272, App
5	10	1.2	206	US-09-252-991A-18034	Sequence 18034, A
6	9	1.1	19	US-09-333-599-3	Sequence 3, Appl1
7	9	1.1	19	US-09-499-781-3	Sequence 3, Appl1
8	8	1.0	71	US-09-369-247-79	Sequence 79, Appl1
9	8	1.0	156	US-09-252-991A-32450	Sequence 32450, A
10	8	1.0	221	US-09-599-360B-104	Sequence 104, App
11	8	1.0	703	US-09-252-991A-17865	Sequence 17865, A
12	7	0.8	57	US-09-082-358B-43	Sequence 43, Appl1
13	7	0.8	66	US-09-205-258-381	Sequence 381, App
14	7	0.8	98	US-09-732-210-341	Sequence 341, App
15	7	0.8	77	US-08-946-329A-94	Sequence 94, Appl1
16	7	0.8	104	US-08-777-708C-5	Sequence 29655, A
17	7	0.8	104	US-09-252-991A-29655	Sequence 4768, App
18	7	0.8	105	US-09-328-352-4768	Sequence 6744, App
19	7	0.8	122	US-09-107-532A-6744	Sequence 6544, App
20	7	0.8	136	US-09-328-352-6544	Sequence 2, Appl1
21	7	0.8	147	US-08-986-837-2	Sequence 18492, A
22	7	0.8	147	US-09-809-593-2	Sequence 32389, A
23	7	0.8	149	US-09-252-991A-18492	Sequence 18109, A
24	7	0.8	155	US-09-252-991A-32399	Sequence 27763, A
25	7	0.8	187	US-09-252-991A-18109	Sequence 20506, A
26	7	0.8	222	US-09-252-991A-27763	
27	7	0.8	273	US-09-252-991A-20506	

28	7	0.8	278	US-09-252-991A-20931	Sequence 20931, A
29	7	0.8	352	US-09-107-532A-6703	Sequence 6703, App
30	7	0.8	355	US-09-134-001C-5391	Sequence 5391, App
31	7	0.8	359	US-09-252-991A-26124	Sequence 26124, A
32	7	0.8	377	US-09-161-994A-12	Sequence 12, Appl1
33	7	0.8	394	US-09-390-721-5	Sequence 5, Appl1
34	7	0.8	434	US-09-252-991A-23489	Sequence 23489, A
35	7	0.8	436	US-09-252-991A-23267	Sequence 23267, A
36	7	0.8	437	US-09-252-991A-24572	Sequence 24572, A
37	7	0.8	452	US-09-328-352-7162	Sequence 7162, App
38	7	0.8	473	US-09-252-991A-28845	Sequence 28845, A
39	7	0.8	493	US-09-117-250-3	Sequence 3, Appl1
40	7	0.8	499	US-09-252-991A-21491	Sequence 21491, A
41	7	0.8	506	US-09-117-250-1	Sequence 9, Appl1
42	7	0.8	531	US-08-933-750C-9	Sequence 9, Appl1
43	7	0.8	531	US-09-234-613-9	Sequence 17757, A
44	7	0.8	534	US-09-252-991A-17757	Sequence 17757, A
45	7	0.8	539	US-09-252-991A-29555	Sequence 29555, A

## ALIGNMENTS

```
RESULT 1
US-09-489-847-273
; Sequence 273, Application US/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
; APPLICANT: Rosen et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031p1
; CURRENT APPLICATION NUMBER: US/09/489,847
; CURRENT FILING DATE: 2000-01-24
; EARLIER APPLICATION NUMBER: PCT/US99/17130
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/096,319
; EARLIER FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: 60/095,454
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/095,455
; EARLIER FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 273
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-489-847-273
```

Query Match 49.4%; Score 413; DB 4; Length 414;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
1 MAGNCGVSIALGVLLGAARLPFGAFAFIALPRESNTIVLKGPTTLAKCYVI 60
1 MAGNCGVSIALGVLLGAARLPFGAFAFIALPRESNTIVLKGPTTLAKCYVI 60
MAGNCGVSIALGVLLGAARLPFGAFAFIALPRESNTIVLKGPTTLAKCYVI 60
SKRHTMTLSGERTVTFSCQSPENHFVLEIQNDICMSGPCFGEVQLOPSTSLPT 120
SKRHTMTLSGERTVTFSCQSPENHFVLEIQNDICMSGPCFGEVQLOPSTSLPT 120
SKRHTMTLSGERTVTFSCQSPENHFVLEIQNDICMSGPCFGEVQLOPSTSLPT 120
LNRFTIWDVKAHKSIGLELOFSIPRLROIGPESCPDVTSHISGRIDATVVRIGTFCSN 180
LNRFTIWDVKAHKSIGLELOFSIPRLROIGPESCPDVTSHISGRIDATVVRIGTFCSN 180
LNRFTIWDVKAHKSIGLELOFSIPRLROIGPESCPDVTSHISGRIDATVVRIGTFCSN 180
GTVSRITKQEGVKALHPFPHPRVSGFSTANRSSIRLCTIESVFGEBSATLMSANY 240
GTVSRITKQEGVKALHPFPHPRVSGFSTANRSSIRLCTIESVFGEBSATLMSANY 240
GTVSRITKQEGVKALHPFPHPRVSGFSTANRSSIRLCTIESVFGEBSATLMSANY 240
```

QY 241 PEGFPEDELTWQFVPAHLRASVSFLNRLNLSNCKEKEVEYYIGSTTNPVEFKLEDK 300  
DB 241 PEGFPEDELTWQFVPAHLRASVSFLNRLNLSNCKEKEVEYYIGSTTNPVEFKLEDK 300  
QY 301 QPQNMAGNPNLSLQGGDDDAQSPGILRLQFOVLVQHPONESNKIYVVDLSNEBAMSLTIE 360  
DB 301 QPQNMAGNPNLSLQGGDDDAQSPGILRLQFOVLVQHPONESNKIYVVDLSNEBAMSLTIE 360  
QY 361 PRPVKQSRKFPVPCFVCLSESRCTSSNLTITSGSKHKISFLCDDLTRLMNMVEX 413  
DB 361 PRPVKQSRKFPVPCFVCLSESRCTSSNLTITSGSKHKISFLCDDLTRLMNMVEX 413

## RESULT 2

US-09-847-271  
; Sequence 271, Application US/09489847  
; Patent No. 6476195  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al  
; TITLE OF INVENTION: 98 Human Secreted Proteins  
; FILE REFERENCE: P2031P1  
; CURRENT APPLICATION NUMBER: US/09/489,847  
; CURRENT FILING DATE: 2000-01-24  
; EARLIER APPLICATION NUMBER: PCT/US99/17130  
; EARLIER FILING DATE: 1999-07-29  
; EARLIER APPLICATION NUMBER: 60/094,657  
; EARLIER FILING DATE: 1998-07-30  
; EARLIER APPLICATION NUMBER: 60/095,486  
; EARLIER FILING DATE: 1998-08-05  
; EARLIER APPLICATION NUMBER: 60/096,319  
; EARLIER FILING DATE: 1998-08-12  
; EARLIER APPLICATION NUMBER: 60/095,454  
; EARLIER FILING DATE: 1998-08-06  
; EARLIER APPLICATION NUMBER: 60/095,455  
; EARLIER FILING DATE: 1998-08-06  
; NUMBER OF SEQ ID NOS: 376  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 271  
; LENGTH: 443  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-489-847-271

Query Match 49.4%; Score 413; DB 4; Length 443;  
Best Local Similarity 100.0%; Pred. No. 0;

Matches 413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGLNCVSIALLGVLLGAARLPRGAFAFEIALPRESNITVLIKGTPTLLAKPCYIYI 60  
DB 30 MAGLNCVSIALLGVLLGAARLPRGAFAFEIALPRESNITVLIKGTPTLLAKPCYIYI 89  
QY 61 SKRHITMLSIKSGERIVFTFSCSPENHFIYEIOKNIDCMGSPGPEYLOPSTSLPT 120  
DB 90 SKRHITMLSIKSGERIVFTFSCSPENHFIYEIOKNIDCMGSPGPEYLOPSTSLPT 149  
QY 121 LNRFTIWDVKAHKSIGLELOFSIPRLRQIGPGBSCPDVYTHISGRIDATVVRIGTFCSN 180  
DB 150 LNRFTIWDVKAHKSIGLELOFSIPRLRQIGPGBSCPDVYTHISGRIDATVVRIGTFCSN 209  
QY 181 GYVSRIKMOEGVMAHLMPFHRNVSQFSIANRSSIKRLCTIESFEEBGSATLMSANY 240  
DB 210 GYVSRIKMOEGVMAHLMPFHRNVSQFSIANRSSIKRLCTIESFEEBGSATLMSANY 269  
QY 241 PEGFPEDELTWQFVPAHLRASVSFLNRLNLSNCKEKEVEYYIGSTTNPVEFKLEDK 300  
DB 270 PEGFPEDELTWQFVPAHLRASVSFLNRLNLSNCKEKEVEYYIGSTTNPVEFKLEDK 329  
QY 301 QPQNMAGNPNLSLQGGDDDAQSPGILRLQFOVLVQHPONESNKIYVVDLSNEBAMSLTIE 360  
DB 330 QPQNMAGNPNLSLQGGDDDAQSPGILRLQFOVLVQHPONESNKIYVVDLSNEBAMSLTIE 389  
QY 361 PRPVKQSRKFPVPCFVCLSESRCTSSNLTITSGSKHKISFLCDDLTRLMNMVEX 413  
DB 361 PRPVKQSRKFPVPCFVCLSESRCTSSNLTITSGSKHKISFLCDDLTRLMNMVEX 413

DB 390 PRPVKQSRKFPVPCFVCLSESRCTSSNLTITSGSKHKISFLCDDLTRLMNMVEX 442

## RESULT 3

US-09-489-847-132  
; Sequence 132, Application US/09489847  
; Patent No. 6476195  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al  
; TITLE OF INVENTION: 98 Human Secreted Proteins  
; FILE REFERENCE: P2031P1  
; CURRENT APPLICATION NUMBER: US/09/489,847  
; CURRENT FILING DATE: 2000-01-24  
; EARLIER APPLICATION NUMBER: PCT/US99/17130  
; EARLIER FILING DATE: 1999-07-29  
; EARLIER APPLICATION NUMBER: 60/094,657  
; EARLIER FILING DATE: 1998-07-30  
; EARLIER APPLICATION NUMBER: 60/095,486  
; EARLIER FILING DATE: 1998-08-05  
; EARLIER APPLICATION NUMBER: 60/096,319  
; EARLIER FILING DATE: 1998-08-12  
; EARLIER APPLICATION NUMBER: 60/095,454  
; EARLIER FILING DATE: 1998-08-06  
; EARLIER APPLICATION NUMBER: 60/095,455  
; EARLIER FILING DATE: 1998-08-06  
; NUMBER OF SEQ ID NOS: 376  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 132  
; LENGTH: 709  
; TYPE: PRT  
; ORGANISM: Homo sapiens

FEATURE: NAME/KEY: SITE

LOCATION: (189)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

FEATURE: NAME/KEY: SITE

LOCATION: (275)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

FEATURE: NAME/KEY: SITE

LOCATION: (414)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

FEATURE: NAME/KEY: SITE

LOCATION: (438)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

FEATURE: NAME/KEY: SITE

LOCATION: (641)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

FEATURE: NAME/KEY: SITE

LOCATION: (643)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

FEATURE: NAME/KEY: SITE

LOCATION: (696)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

FEATURE: NAME/KEY: SITE

LOCATION: (697)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

US-09-489-847-132

Query Match 25.2%; Score 211; DB 4; Length 709;  
Best Local Similarity 99.5%; Pred. No. 2.2e-203;

Matches 411; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAGLNCVSIALLGVLLGAARLPRGAFAFEIALPRESNITVLIKGTPTLLAKPCYIYI 60  
DB 1 MAGLNCVSIALLGVLLGAARLPRGAFAFEIALPRESNITVLIKGTPTLLAKPCYIYI 60

```
OY 61 SKRHTMTLSKSGERIVTFPSQSPENHFVIEIQNIDCMGSPCFGEVOLQPSSTLLPT 120
DB 61 SKRHTMTLSKSGERIVTFPSQSPENHFVIEIQNIDCMGSPCFGEVOLQPSSTLLPT 120
OY 121 LNRFTFMVKAHKSIGLELOFSIPRLROIIPGSCSPDGVTHSISGRIDATVVRIGTFCSN 180
DB 121 LNRFTFMVKAHKSIGLELOFSIPRLROIIPGSCSPDGVTHSISGRIDATVVRIGTFCSN 180
OY 181 GTVSRKMQEGVKALHLPWFPRNVSGFSIANRSSIKRLCIIESVFEGEGSATLSMANY 240
DB 181 GTVSRKMQEGVKALHLPWFPRNVSGFSIANRSSIKRLCIIESVFEGEGSATLSMANY 240
OY 241 PEGFPEDELMTQFVVPAPHLRASVSFLNPNUNSCERKEBEVEYYPGSTTNPVEVKLEDK 300
DB 241 PEGFPEDELMTQFVVPAPHLRASVSFLNPNUNSCERKEBEVEYYPGSTTNPVEVKLEDK 300
OY 301 QPGNAGNPNLSLQSCDDQSPGILRLQFVLVQHPQNSKITYVVLSEKRAVSLTIE 360
DB 301 QPGNAGNPNLSLQSCDDQSPGILRLQFVLVQHPQNSKITYVVLSEKRAVSLTIE 360
OY 361 PRPVQSKRFVPGCFVCLSESRSCSNLTLTSGSKHKISFLCDLTRLMMNVEK 413
DB 361 PRPVQSKRFVPGCFVCLSESRSCSNLTLTSGSKHKISFLCDLTRLMMNVEK 413
```

## RESULT 4

```
US-09-489-847-272
; Sequence 272, Application US/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
; APPLICANT: Rosen et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031P1
; CURRENT APPLICATION NUMBER: US/09/489,847
; CURRENT FILING DATE: 2000-01-24
; EARLIER APPLICATION NUMBER: PCT/US99/17130
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/096,319
; EARLIER FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: 60/095,454
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/095,455
; EARLIER FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 272
; LENGTH: 159
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-489-847-272
```

```
Query Match 19.0%; Score 159; DB 4; Length 159;
Best Local Similarity 100.0%; Pred. No. 1.1e-151;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 30 FEIALPRESNITVLKLTGPTLLAKPCYIVISKRHITMISKGRIVTFPSQSPENHF 89
DB 1 FEIALPRESNITVLKLTGPTLLAKPCYIVISKRHITMISKGRIVTFPSQSPENHF 60
OY 90 VIEIQNIDCMGSPCFGEVOLQPSSTLLPTLNRFTFMVKAHKSIGLELOFSIRLROI 149
DB 61 VIEIQNIDCMGSPCFGEVOLQPSSTLLPTLNRFTFMVKAHKSIGLELOFSIRLROI 120
OY 150 GPGSCPDGVTHSISGRIDATVVRIGTFCSNGTVRIMK 188
DB 121 GPGSCPDGVTHSISGRIDATVVRIGTFCSNGTVRIMK 159
```

## RESULT 5

```
US-09-252-991A-18034
; Sequence 18034, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18034
; LENGTH: 206
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18034
```

```
Query Match 1.2%; Score 10; DB 4; Length 206;
Best Local Similarity 100.0%; Pred. No. 0.099;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 19 GAARLPFGAE 28
DB 135 GAARLPFGAE 144
```

## RESULT 6

```
US-09-333-599-3
; Sequence 3, Application US/09333599
; Patent No. 6245898
; GENERAL INFORMATION:
; APPLICANT: Testa, Jacqueline E.
; APPLICANT: Quigley, James P.
; TITLE OF INVENTION: MONOCLONAL ANTIBODIES THAT RECOGNIZE ANTIGENS
; TITLE OF INVENTION: ASSOCIATED WITH TUMOR METASTASIS
; FILE REFERENCE: SUNY
; CURRENT APPLICATION NUMBER: US/09/333,599
; CURRENT FILING DATE: 1999-06-15
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 19
; TYPE: PRT
; ORGANISM: synthetic construct
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (15)
; OTHER INFORMATION: Xaa is Gln or Ile.
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (17)
US-09-333-599-3
```

```
Query Match 1.1%; Score 9; DB 3; Length 19;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 30 FEIALPRES 38
DB 1 FEIALPRES 9
```

```
RESULT 7
US-09-499-781-3
; Sequence 3, Application US/09499781
; Patent No. 6498014
; GENERAL INFORMATION:
; APPLICANT: Testa, Jacqueline E.
```

APPLICANT: Quigley, James P.  
APPLICANT: Seandel, Marco  
TITLE OF INVENTION: MONOCLONAL ANTIBODIES THAT RECOGNIZE ANTIGENS  
TITLE OF INVENTION: ASSOCIATED WITH TUMOR METASTASIS  
FILE REFERENCE: SUNY  
CURRENT APPLICATION NUMBER: US/09/499,781  
CURRENT FILING DATE: 2000-02-08  
PRIOR APPLICATION NUMBER: 09/333,599  
PRIOR FILING DATE: 1999-06-15  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 3  
LENGTH: 19  
TYPE: PRT  
ORGANISM: synthetic construct  
FEATURE:  
NAME/KEY: UNSURE  
LOCATION: (15)  
OTHER INFORMATION: Xaa is Gln or Ile.  
NAME/KEY: UNSURE  
LOCATION: (17)  
US-09-499-781-3

Query Match 1.0%; Score 9; DB 4; Length 19;  
Best Local Similarity 100.0%; Pred. No. 0.11;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 FEIAPRES 38  
DB 1 FEIAPRES 9

RESULT 8  
US-09-369-247-79  
Sequence 79, Application US/09369247  
Patent No. 6569992  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: 44 Human Secreted Proteins  
FILE REFERENCE: P2024P1  
CURRENT APPLICATION NUMBER: US/09/369,247  
CURRENT FILING DATE: 1999-08-05  
EARLIER APPLICATION NUMBER: 60/074,118  
EARLIER FILING DATE: 1998-02-09  
EARLIER APPLICATION NUMBER: 60/074,157  
EARLIER FILING DATE: 1998-02-09  
EARLIER APPLICATION NUMBER: 60/074,137  
EARLIER FILING DATE: 1998-02-09  
EARLIER APPLICATION NUMBER: 60/074,341  
EARLIER FILING DATE: 1998-02-09  
EARLIER APPLICATION NUMBER: 60/074,141  
EARLIER FILING DATE: 1998-02-09  
NUMBER OF SEQ ID NOS: 172  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 79  
LENGTH: 71  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (71)  
OTHER INFORMATION: Xaa equals stop translation  
US-09-369-247-79

Query Match 1.0%; Score 8; DB 4; Length 71;  
Best Local Similarity 100.0%; Pred. No. 3.8;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AGIACGVS 9  
DB 41 AGIACGVS 48

RESULT 9  
US-09-252-991A-32450  
Sequence 32450, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 32450  
LENGTH: 156  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-32450

Query Match 1.0%; Score 8; DB 4; Length 156;  
Best Local Similarity 100.0%; Pred. No. 7.8;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GAARLPRG 26  
DB 32 GAARLPRG 39

RESULT 10  
US-09-599-360B-104  
Sequence 104, Application US/09599360B  
Patent No. 6548633  
GENERAL INFORMATION:  
APPLICANT: Dumas Milne Edwards, J.B.  
APPLICANT: Bouquelerec, L.  
TITLE OF INVENTION: Complementary DNA's Encoding Proteins with Signal Peptides  
FILE REFERENCE: GENSET.050CP3  
CURRENT APPLICATION NUMBER: US/09/599,360B  
CURRENT FILING DATE: 2000-06-21  
PRIOR APPLICATION NUMBER: 60/113,686  
PRIOR FILING DATE: 1998-12-22  
PRIOR APPLICATION NUMBER: 60/141,032  
PRIOR FILING DATE: 1999-06-25  
PRIOR APPLICATION NUMBER: 09/469,099  
PRIOR FILING DATE: 1999-12-21  
NUMBER OF SEQ ID NOS: 123  
SOFTWARE: Patent.pm  
SEQ ID NO 104  
LENGTH: 221  
TYPE: PRT  
ORGANISM: Homo Sapiens  
FEATURE:  
NAME/KEY: SIGNAL  
LOCATION: -28...-1  
US-09-599-360B-104

Query Match 1.0%; Score 8; DB 4; Length 221;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 LIGVLLLG 19  
DB 104 LIGVLLLG 111

RESULT 11  
US-09-252-991A-17865  
Sequence 17865, Application US/09252991A  
Patent No. 6551795

GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 17865  
LENGTH: 703  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-17865

Query Match 1.0%; Score 8; DB 4; Length 703;  
Best Local Similarity 100.0%; Pred. No. 32;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 11 ALLGVLL 18  
Db 386 ALLGVLL 393

RESULT 12  
US-09-082-358B-43  
Sequence 43, Application US/09082358B  
Patent No. 6469153  
GENERAL INFORMATION:  
APPLICANT: Goff, Stephen P.  
TITLE OF INVENTION: EIP-1, EIP-3 GENES, ENVELOPE-INTERACTING PROTEINS,  
FILE REFERENCE: 0575/54804  
CURRENT APPLICATION NUMBER: US/09/082,358B  
CURRENT FILING DATE: 1998-05-20  
NUMBER OF SEQ ID NOS: 106  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 43  
LENGTH: 57  
TYPE: PRT  
ORGANISM: murine  
US-09-082-358B-43

Query Match 0.8%; Score 7; DB 4; Length 57;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 692 KKKKKKT 698  
Db 50 KKKKKKT 56

RESULT 13  
US-09-205-258-381  
Sequence 381, Application US/09205258  
Patent No. 6525174  
GENERAL INFORMATION:  
APPLICANT: Young et al.  
TITLE OF INVENTION: 207 Human Secreted Proteins  
FILE REFERENCE: P200781  
CURRENT APPLICATION NUMBER: US/09/205,258  
CURRENT FILING DATE: 1998-12-04  
EARLIER APPLICATION NUMBER: PCT/US98/11422  
EARLIER FILING DATE: 1998-06-04  
EARLIER APPLICATION NUMBER: 60/048,885  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,375  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,881

EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,880  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,896  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,020  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,876  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,895  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,884  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,894  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,971  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,964  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,882  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,899  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,901  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,892  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,915  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,019  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,970  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,972  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,916  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,373  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,875  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,374  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,917  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,949  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,974  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,883  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,897  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,898  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,962  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,963  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,877  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,878  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/070,923  
EARLIER FILING DATE: 1997-12-18  
EARLIER APPLICATION NUMBER: 60/092,921  
EARLIER FILING DATE: 1998-07-15  
EARLIER APPLICATION NUMBER: 60/094,657  
EARLIER FILING DATE: 1998-07-30

NUMBER OF SEQ ID NOS: 1227  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 381  
LENGTH: 66  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (14)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (62)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-205-258-381

Query Match 0.8%; Score 7; DB 4; Length 66;  
Best Local Similarity 100.0%; Pred. No. 36;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 692 KKKKKKT 698  
DB 49 KKKKKKT 55

RESULT 14  
US-09-732-210-341  
Sequence 341, Application US/09732210  
Patent No. 6573361  
GENERAL INFORMATION:  
APPLICANT: Bunkers, Greg J.  
APPLICANT: Liang, Jihong  
APPLICANT: Mitcank, Cindy A.  
APPLICANT: Seale, Jeffrey W.  
APPLICANT: Wu, Yonnie S.  
TITLE OF INVENTION: Anti-fungal Proteins and Methods for Their Use  
FILE REFERENCE: 38-21 (15036) B  
CURRENT APPLICATION NUMBER: US/09/732,210  
CURRENT FILING DATE: 2000-12-07  
PRIOR APPLICATION NUMBER: US 60/169,513  
PRIOR FILING DATE: 1999-12-07  
PRIOR APPLICATION NUMBER: US 60/169,340  
PRIOR FILING DATE: 1999-12-07  
NUMBER OF SEQ ID NOS: 1753  
SEQ ID NO 341  
LENGTH: 78  
TYPE: PRT  
ORGANISM: Dictyostelium discoideum  
US-09-732-210-341

Query Match 0.8%; Score 7; DB 4; Length 78;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 692 KKKKKKT 698  
DB 3 KKKKKKT 9

RESULT 15  
US-08-946-329A-94  
Sequence 94, Application US/08946329A  
Patent No. 6057091  
GENERAL INFORMATION:  
APPLICANT: Beachy, Philip A.  
APPLICANT: Porter, Jeffrey A.  
TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES  
NUMBER OF SEQUENCES: 109  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 4225 Executive Square, Suite 1400  
CITY: La Jolla  
STATE: CA

COUNTRY: USA  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows 95  
SOFTWARE: FASTSEQ for Windows Version 2.0b  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/946,329A  
FILING DATE: 07-OCT-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/061,323  
FILING DATE: 07-OCT-1996  
APPLICATION NUMBER: 08/729,743  
FILING DATE: 10-JUL-1996  
APPLICATION NUMBER: 08/567,357  
FILING DATE: 04-DEC-1995  
APPLICATION NUMBER: 08/349,498  
FILING DATE: 02-DEC-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Haile, Lisa A.  
REGISTRATION NUMBER: 38,347  
REFERENCE/DOCKET NUMBER: 07265/140001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619/678-5070  
TELEFAX: 619/678-5099  
INFORMATION FOR SEQ ID NO: 94:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 97 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-946-329A-94

Query Match 0.8%; Score 7; DB 3; Length 97;  
Best Local Similarity 100.0%; Pred. No. 51;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 615 AEEIFSL 621  
DB 17 AEEIFSL 23

Search completed: February 20, 2004, 19:51:00  
Job time : 33 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2004, 19:46:38 ; Search time 100 Seconds  
(without alignments)  
1750.436 Million cell updates/sec

Title: US-09-899-569a-4

Perfect score: 836  
Sequence: 1 MAGLNCGVSIALLGVLLGA.....SSKTDIPULNTQEPMEPAE 836

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 801455 seqs, 209382283 residues

Word size : 0

Total number of hits satisfying chosen parameters: 801455

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	836	100.0	836	US-09-899-569a-4	Sequence 4, App11
2	650	77.8	749	US-09-899-569a-2	Sequence 2, App11
3	341	40.8	343	US-10-137-870-162	Sequence 162, App
4	341	40.8	343	US-10-140-018-162	Sequence 162, App
5	341	40.8	343	US-10-140-021-162	Sequence 162, App
6	341	40.8	343	US-10-140-274-162	Sequence 162, App
7	341	40.8	343	US-10-140-471-162	Sequence 162, App
8	341	40.8	343	US-10-140-807-162	Sequence 162, App
9	341	40.8	343	US-10-140-922-162	Sequence 162, App
10	341	40.8	343	US-10-140-924-162	Sequence 162, App
11	341	40.8	343	US-10-140-926-162	Sequence 162, App
12	341	40.8	343	US-10-141-698-162	Sequence 162, App
13	341	40.8	343	US-10-141-702-162	Sequence 162, App
14	341	40.8	343	US-10-141-704-162	Sequence 162, App
15	341	40.8	343	US-10-142-421-162	Sequence 162, App

16	341	40.8	343	12	US-10-142-432-162	Sequence 162, App
17	341	40.8	343	12	US-10-142-767-162	Sequence 162, App
18	341	40.8	343	12	US-10-143-033-162	Sequence 162, App
19	341	40.8	343	12	US-10-144-994-162	Sequence 162, App
20	341	40.8	343	12	US-10-145-628-162	Sequence 162, App
21	341	40.8	343	12	US-10-145-631-162	Sequence 162, App
22	341	40.8	343	12	US-10-145-633-162	Sequence 162, App
23	341	40.8	343	12	US-10-145-746-162	Sequence 162, App
24	341	40.8	343	12	US-10-145-748-162	Sequence 162, App
25	341	40.8	343	12	US-10-145-823-162	Sequence 162, App
26	341	40.8	343	12	US-10-145-826-162	Sequence 162, App
27	341	40.8	343	12	US-10-145-870-162	Sequence 162, App
28	341	40.8	343	12	US-10-145-876-162	Sequence 162, App
29	341	40.8	343	12	US-10-145-959-162	Sequence 162, App
30	341	40.8	343	12	US-10-146-724-162	Sequence 162, App
31	341	40.8	343	12	US-10-146-725-162	Sequence 162, App
32	341	40.8	343	12	US-10-146-795-162	Sequence 162, App
33	341	40.8	343	12	US-10-147-495-162	Sequence 162, App
34	341	40.8	343	12	US-10-147-501-162	Sequence 162, App
35	341	40.8	343	12	US-10-147-504-162	Sequence 162, App
36	341	40.8	343	12	US-10-147-509-162	Sequence 162, App
37	341	40.8	343	12	US-10-147-510-162	Sequence 162, App
38	341	40.8	343	12	US-10-147-511-162	Sequence 162, App
39	341	40.8	343	12	US-10-147-529-162	Sequence 162, App
40	341	40.8	343	12	US-10-152-397-162	Sequence 162, App
41	341	40.8	343	12	US-10-153-586-162	Sequence 162, App
42	341	40.8	343	12	US-10-158-793-162	Sequence 162, App
43	341	40.8	343	12	US-10-158-786-162	Sequence 162, App
44	341	40.8	343	12	US-10-140-019-162	Sequence 162, App
45	341	40.8	343	12	US-10-140-019-162	Sequence 162, App

#### ALIGNMENTS

US-09-899-569a-4

Sequence 4, Application US/09899569A

Patent No. US20020142003A1

GENERAL INFORMATION:

APPLICANT: No. US20020142003Albert Schweitzer

APPLICANT: Marwa Scherl-Mostageer

APPLICANT: Wolfgang Sommergruber

APPLICANT: Roger Absseher

TITLE OF INVENTION: Tumorsozioleites Antigen (B345)

FILE REFERENCE: 0652.2280001

CURRENT APPLICATION NUMBER: US/09/899,569A

CURRENT FILING DATE: 2001-07-06

PRIOR FILING DATE: 2000-07-07

PRIOR APPLICATION NUMBER: DE 100 33 080.0

PRIOR FILING DATE: 2001-04-19

PRIOR APPLICATION NUMBER: US 60/243,158

PRIOR FILING DATE: 2000-10-25

PRIOR APPLICATION NUMBER: US 60/297,747

PRIOR FILING DATE: 2001-06-14

NUMBER OF SEQ ID NOS: 40

SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 4

LENGTH: 836

TYPE: PRT

ORGANISM: Homo sapiens

US-09-899-569a-4

Query Match 100.0%; Score 836; DB 10; Length 836;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 836; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGLNCGVSIALLGVLLGAARLPARGAFAFALPRESNITVLTGPTTLAKCYIVT 60

DB 1 MAGLNCGVSIALLGVLLGAARLPARGAFAFALPRESNITVLTGPTTLAKCYIVT 60

QY 61 SKRHITMLSIKSGRIIVTFSCQSPENHFVIEIQNIDMSGPCPFGEVQLQPSLPT 120

```

Db      61 SKRHITMLSIKSGERIVFTFSCSPENHFWIEIQKNIDCMGCPGEGVQLPSTSLLP 120
Qy      121 LNRTFMDVKAHKSIGLELOFSIRLROIGPGECPDGVTHSISGRIDATVVRIGTFCSN 180
Db      121 LNRTFMDVKAHKSIGLELOFSIRLROIGPGECPDGVTHSISGRIDATVVRIGTFCSN 180
Qy      181 GVSIRIKMOEGVMALHLPWFHPRNVSGFSIANRSSIKRLCIIESVFESEGSATLMSANY 240
Db      181 GVSIRIKMOEGVMALHLPWFHPRNVSGFSIANRSSIKRLCIIESVFESEGSATLMSANY 240
Qy      241 PGGFPEDELTMOQFVPAHLRASVSFLNPNLSNCEKKEEREVEYITPGSTTNPBEVFLDEK 300
Db      241 PGGFPEDELTMOQFVPAHLRASVSFLNPNLSNCEKKEEREVEYITPGSTTNPBEVFLDEK 300
Qy      301 QGNGMAGNNTLSQCGDDAOSPGILRLQFOVLVOPONESNKIYVDLSNERAMSLTIE 360
Db      301 QGNGMAGNNTLSQCGDDAOSPGILRLQFOVLVOPONESNKIYVDLSNERAMSLTIE 360
Qy      361 PRPVKSRKFPVPCFVLESRTCSNLTJSGSKHISFLCDDLTRLMMNVEKTIISCTDH 420
Db      361 PRPVKSRKFPVPCFVLESRTCSNLTJSGSKHISFLCDDLTRLMMNVEKTIISCTDH 420
Qy      421 RVCQRKSYSLQVPSDILHLPVELHDFSWMKLVPRKLSLVPAQKLOQTHEKPCNTSF 480
Db      421 RVCQRKSYSLQVPSDILHLPVELHDFSWMKLVPRKLSLVPAQKLOQTHEKPCNTSF 480
Qy      481 SYLVASAIPODLYFSGFCPGGSIKOIOVKONISVTLRTFAPSFOEASROGLTVSFIFY 540
Db      481 SYLVASAIPODLYFSGFCPGGSIKOIOVKONISVTLRTFAPSFOEASROGLTVSFIFY 540
Qy      541 FKEEGVFTVPTDKSKVYLRTPNMDRGLPSLTSVSNISVPRDQVACLTFPKERSGVCC 600
Db      541 FKEEGVFTVPTDKSKVYLRTPNMDRGLPSLTSVSNISVPRDQVACLTFPKERSGVCC 600
Qy      601 TGRAFMTIOEORTRAEISLDEVDLPKPSFHHHSFWNINISNCSPTSGKOLDLFSVTLT 660
Db      601 TGRAFMTIOEORTRAEISLDEVDLPKPSFHHHSFWNINISNCSPTSGKOLDLFSVTLT 660
Qy      661 PRVVDLTVLILAAVGGVLLLSALGLIICVKKKKKTKNKPAGIYNGNINTEMROPK 720
Db      661 PRVVDLTVLILAAVGGVLLLSALGLIICVKKKKKTKNKPAGIYNGNINTEMROPK 720
Qy      721 KFOGKRDNDSHYAVIEDTMVYGLLQDSGSLQPEVDYTRPFQGTGVCPSPTIC 780
Db      721 KFOGKRDNDSHYAVIEDTMVYGLLQDSGSLQPEVDYTRPFQGTGVCPSPTIC 780
Qy      781 SRAPTAKLATEEPSPSPSESEPTTFSHPNNGDVSSKOTDILPLNTQEPMEPAR 836
Db      781 SRAPTAKLATEEPSPSPSESEPTTFSHPNNGDVSSKOTDILPLNTQEPMEPAR 836

```

```

RESULT 2
US-09-899-569a-2
; Sequence 2, Application US/09899569A
; Patent No. US20020142003A1
; GENERAL INFORMATION:
; APPLICANT: No. US20020142003Albert Schweitzer
; APPLICANT: Marwa Scherl-Mostageer
; APPLICANT: Wolfgang Sommergruber
; APPLICANT: Roger Abseher
; TITLE OF INVENTION: Tumoraesoziiertes Antigen (B345)
; FILE REFERENCE: 0652.2280001
; CURRENT APPLICATION NUMBER: US/09/899,569A
; PRIOR FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: DE 100 33 080.0
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: DE 101 19 294.0
; PRIOR FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: US 60/243,158
; PRIOR FILING DATE: 2000-10-25
; PRIOR APPLICATION NUMBER: US 60/297,747
; PRIOR FILING DATE: 2001-06-14

```

```

; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 749
; TYPE: prt
; ORGANISM: Homo sapiens
US-09-899-569a-2
Query Match      77.8%; Score 650; DB 10; Length 749;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 650; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      67 MLSISGERIVFTFSCSPENHFWIEIQKNIDCMGCPGEGVQLPSTSLLP 126
Db      1 MLSISGERIVFTFSCSPENHFWIEIQKNIDCMGCPGEGVQLPSTSLLP 126
Qy      127 WPKAKHKSIGLELOFSIRLROIGPGECPDGVTHSISGRIDATVVRIGTFCSN 186
Db      61 WPKAKHKSIGLELOFSIRLROIGPGECPDGVTHSISGRIDATVVRIGTFCSN 186
Qy      187 KMOEGVMALHLPWFHPRNVSGFSIANRSSIKRLCIIESVFESEGSATLMSANYEGPPE 246
Db      121 KMOEGVMALHLPWFHPRNVSGFSIANRSSIKRLCIIESVFESEGSATLMSANYEGPPE 180
Qy      247 DELMTQFVPAHLRASVSFLNPNLSNCEKKEEREVEYITPGSTTNPBEVFLDEK 306
Db      181 DELMTQFVPAHLRASVSFLNPNLSNCEKKEEREVEYITPGSTTNPBEVFLDEK 306
Qy      307 GNFNLSQCGDDAOSPGILRLQFOVLVOPONESNKIYVDLSNERAMSLTIEPRPVK 366
Db      241 GNFNLSQCGDDAOSPGILRLQFOVLVOPONESNKIYVDLSNERAMSLTIEPRPVK 300
Qy      367 SRKFVPGFCVLESRTCSNLTJSGSKHISFLCDDLTRLMMNVEKTIISCTDHRVCQRK 426
Db      301 SRKFVPGFCVLESRTCSNLTJSGSKHISFLCDDLTRLMMNVEKTIISCTDHRVCQRK 360
Qy      427 SYSLQVPSDILHLPVELHDFSWMKLVPRKLSLVPAQKLOQTHEKPCNTSF 486
Db      361 SYSLQVPSDILHLPVELHDFSWMKLVPRKLSLVPAQKLOQTHEKPCNTSF 420
Qy      487 AIPSDOLYFSGFCPGGSIKOIOVKONISVTLRTFAPSFOEASROGLTVSFIFY 546
Db      421 AIPSDOLYFSGFCPGGSIKOIOVKONISVTLRTFAPSFOEASROGLTVSFIFY 480
Qy      547 FTVTPTDKSKVYLRTPNMDRGLPSLTSVSNISVPRDQVACLTFPKERSGVCC 606
Db      481 FTVTPTDKSKVYLRTPNMDRGLPSLTSVSNISVPRDQVACLTFPKERSGVCC 540
Qy      607 IIOEORTRAEISLDEVDLPKPSFHHHSFWNINISNCSPTSGKOLDLFSVTLT 666
Db      541 IIOEORTRAEISLDEVDLPKPSFHHHSFWNINISNCSPTSGKOLDLFSVTLT 600
Qy      667 TVILIAAVGGVLLLSALGLIICVKKKKKTKNKPAGIYNGNINTEM 716
Db      601 TVILIAAVGGVLLLSALGLIICVKKKKKTKNKPAGIYNGNINTEM 650

```

```

RESULT 3
US-10-137-870-162
; Sequence 162, Application US/10137870
; Publication No. US20030138883A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven

```



APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3330R1C155  
CURRENT APPLICATION NUMBER: US/10/137,870  
CURRENT FILING DATE: 2002-05-03  
Prior Application removed - See Palm or File Wrapper  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 162  
LENGTH: 343  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-10-137-870-162

Query Match 40.8%; Score 341; DB 12; Length 343;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGLNGVSIALGVLLGAARLPGAAFEIALPRESNITVLKGLPTLLAKPCYIV 60  
DB 1 MAGLNGVSIALGVLLGAARLPGAAFEIALPRESNITVLKGLPTLLAKPCYIV 60  
QY 61 SKRHITMISGGERIVTFSCSPENHFVIEIQNIDMSGPCPGEVQLPSTSLPT 120  
DB 61 SKRHITMISGGERIVTFSCSPENHFVIEIQNIDMSGPCPGEVQLPSTSLPT 120  
QY 121 LNRFTIMDVKAHKSIGLELQFSIPRLROI GPESCPDGVTHSISGRIDATVVRIGTFCSN 180  
DB 121 LNRFTIMDVKAHKSIGLELQFSIPRLROI GPESCPDGVTHSISGRIDATVVRIGTFCSN 180  
QY 181 GTVSRITKQEGVKMALHPWFHPRVSGFSIANRSSIKRLCTIESVFEGEGSATLMSANY 240  
DB 181 GTVSRITKQEGVKMALHPWFHPRVSGFSIANRSSIKRLCTIESVFEGEGSATLMSANY 240  
QY 241 PEGPREDLMTQFVVPALRASVSFLNPNLNCERKEERYIIPGSTTNEVEVKLEDK 300  
DB 241 PEGPREDLMTQFVVPALRASVSFLNPNLNCERKEERYIIPGSTTNEVEVKLEDK 300  
QY 301 QPGNMGNFNLSLQSCDDQDQSPGILRLQFVLVQHPONES 341  
DB 301 QPGNMGNFNLSLQSCDDQDQSPGILRLQFVLVQHPONES 341

## RESULT 4

US-10-140-018-162  
Sequence 162, Application US/10140018  
Publication No. US20030138885A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: DeForge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3330R1C158

CURRENT APPLICATION NUMBER: US/10/140,018  
CURRENT FILING DATE: 2002-05-06  
Prior Application removed - See Palm or File Wrapper  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 162  
LENGTH: 343  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-10-140-018-162

Query Match 40.8%; Score 341; DB 12; Length 343;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGLNGVSIALGVLLGAARLPGAAFEIALPRESNITVLKGLPTLLAKPCYIV 60  
DB 1 MAGLNGVSIALGVLLGAARLPGAAFEIALPRESNITVLKGLPTLLAKPCYIV 60  
QY 61 SKRHITMISGGERIVTFSCSPENHFVIEIQNIDMSGPCPGEVQLPSTSLPT 120  
DB 61 SKRHITMISGGERIVTFSCSPENHFVIEIQNIDMSGPCPGEVQLPSTSLPT 120  
QY 121 LNRFTIMDVKAHKSIGLELQFSIPRLROI GPESCPDGVTHSISGRIDATVVRIGTFCSN 180  
DB 121 LNRFTIMDVKAHKSIGLELQFSIPRLROI GPESCPDGVTHSISGRIDATVVRIGTFCSN 180  
QY 181 GTVSRITKQEGVKMALHPWFHPRVSGFSIANRSSIKRLCTIESVFEGEGSATLMSANY 240  
DB 181 GTVSRITKQEGVKMALHPWFHPRVSGFSIANRSSIKRLCTIESVFEGEGSATLMSANY 240  
QY 241 PEGPREDLMTQFVVPALRASVSFLNPNLNCERKEERYIIPGSTTNEVEVKLEDK 300  
DB 241 PEGPREDLMTQFVVPALRASVSFLNPNLNCERKEERYIIPGSTTNEVEVKLEDK 300  
QY 301 QPGNMGNFNLSLQSCDDQDQSPGILRLQFVLVQHPONES 341  
DB 301 QPGNMGNFNLSLQSCDDQDQSPGILRLQFVLVQHPONES 341

## RESULT 5

US-10-140-021-162  
Sequence 162, Application US/10140021  
Publication No. US20030138886A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: DeForge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3330R1C167  
CURRENT APPLICATION NUMBER: US/10/140,021  
CURRENT FILING DATE: 2002-05-06  
Prior Application removed - See Palm or File Wrapper  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 162  
LENGTH: 343  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-10-140-021-162

Query Match 40.8%; Score 341; DB 12; Length 343;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MAGLNGVSIALGVLLGAARLPAGAFAETALPRESNITVLIKGTPTLLAKPCYIYI 60
DB 1 MAGLNGVSIALGVLLGAARLPAGAFAETALPRESNITVLIKGTPTLLAKPCYIYI 60
QY 61 SKRHITMLSIKSGERIVTFSCQSPENHFVIEIQKNDICMSGPCPFGEVOLQPSSTLLPT 120
DB 61 SKRHITMLSIKSGERIVTFSCQSPENHFVIEIQKNDICMSGPCPFGEVOLQPSSTLLPT 120
QY 121 LNRFTIMVVKAKHSIGLELOFSIPRLROIIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
DB 121 LNRFTIMVVKAKHSIGLELOFSIPRLROIIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
QY 181 GTVSRIRKQEGVKMALHPWFHPRNVSGFSIANRSSIKRLCTIEESVFEGEGSATLMSANY 240
DB 181 GTVSRIRKQEGVKMALHPWFHPRNVSGFSIANRSSIKRLCTIEESVFEGEGSATLMSANY 240
QY 241 PEGPPEDELMTWQFVVPALHRAVSFLNPNISNCRKEKREVEYYIPGSTTNEVFKLEBK 300
DB 241 PEGPPEDELMTWQFVVPALHRAVSFLNPNISNCRKEKREVEYYIPGSTTNEVFKLEBK 300
QY 301 QPGNAGNFNLSLQGCDDQDQSPGILRLQFOVLVQHPONES 341
DB 301 QPGNAGNFNLSLQGCDDQDQSPGILRLQFOVLVQHPONES 341
```

RESULT 6  
US-10-140-274-162  
; Sequence 162, Application US/10140274  
; Publication No. US20030143674A1

GENERAL INFORMATION:

```
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gueney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C161
CURRENT APPLICATION NUMBER: US/10/140,274
CURRENT FILING DATE: 2002-05-06
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 162
LENGTH: 343
TYPE: PRT
ORGANISM: Homo Sapien
US-10-140-274-162
```

Query Match 40.8%; Score 341; DB 12; Length 343;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MAGLNGVSIALGVLLGAARLPAGAFAETALPRESNITVLIKGTPTLLAKPCYIYI 60
DB 1 MAGLNGVSIALGVLLGAARLPAGAFAETALPRESNITVLIKGTPTLLAKPCYIYI 60
```

```
QY 61 SKRHITMLSIKSGERIVTFSCQSPENHFVIEIQKNDICMSGPCPFGEVOLQPSSTLLPT 120
DB 61 SKRHITMLSIKSGERIVTFSCQSPENHFVIEIQKNDICMSGPCPFGEVOLQPSSTLLPT 120
QY 121 LNRFTIMVVKAKHSIGLELOFSIPRLROIIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
DB 121 LNRFTIMVVKAKHSIGLELOFSIPRLROIIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
QY 181 GTVSRIRKQEGVKMALHPWFHPRNVSGFSIANRSSIKRLCTIEESVFEGEGSATLMSANY 240
DB 181 GTVSRIRKQEGVKMALHPWFHPRNVSGFSIANRSSIKRLCTIEESVFEGEGSATLMSANY 240
QY 241 PEGPPEDELMTWQFVVPALHRAVSFLNPNISNCRKEKREVEYYIPGSTTNEVFKLEBK 300
DB 241 PEGPPEDELMTWQFVVPALHRAVSFLNPNISNCRKEKREVEYYIPGSTTNEVFKLEBK 300
QY 301 QPGNAGNFNLSLQGCDDQDQSPGILRLQFOVLVQHPONES 341
DB 301 QPGNAGNFNLSLQGCDDQDQSPGILRLQFOVLVQHPONES 341
```

RESULT 7  
US-10-140-471-162  
; Sequence 162, Application US/10140471  
; Publication No. US2003013887A1

GENERAL INFORMATION:

```
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gueney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C163
CURRENT APPLICATION NUMBER: US/10/140,471
CURRENT FILING DATE: 2002-05-06
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 162
LENGTH: 343
TYPE: PRT
ORGANISM: Homo Sapien
US-10-140-471-162
```

Query Match 40.8%; Score 341; DB 12; Length 343;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MAGLNGVSIALGVLLGAARLPAGAFAETALPRESNITVLIKGTPTLLAKPCYIYI 60
DB 1 MAGLNGVSIALGVLLGAARLPAGAFAETALPRESNITVLIKGTPTLLAKPCYIYI 60
QY 61 SKRHITMLSIKSGERIVTFSCQSPENHFVIEIQKNDICMSGPCPFGEVOLQPSSTLLPT 120
DB 61 SKRHITMLSIKSGERIVTFSCQSPENHFVIEIQKNDICMSGPCPFGEVOLQPSSTLLPT 120
QY 121 LNRFTIMVVKAKHSIGLELOFSIPRLROIIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
DB 121 LNRFTIMVVKAKHSIGLELOFSIPRLROIIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
QY 181 GTVSRIRKQEGVKMALHPWFHPRNVSGFSIANRSSIKRLCTIEESVFEGEGSATLMSANY 240
```

Db 181 GTVSRIRKQEGVKALHLPWFHPRNVSGFSIANSSIRLCIIISVFEGESATLMSANY 240  
QY 241 PEGPPEDELMTWQFVVPALHRAVSFLNFNLSNCRKEERYEYIPGSTTNPVEVKLEDK 300-  
Db 241 PEGPPEDELMTWQFVVPALHRAVSFLNFNLSNCRKEERYEYIPGSTTNPVEVKLEDK 300  
QY 301 QPGMAGNFNLSLOGCDODASPGILRLQFOVLVOHPONES 341  
Db 301 QPGMAGNFNLSLOGCDODASPGILRLQFOVLVOHPONES 341

RESULT 8  
US-10-140-807-162  
; Sequence 162, Application US/10140807  
; Publication No. US20030134354A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: Deforge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Macanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P33081C174  
; CURRENT APPLICATION NUMBER: US/10/140,807  
; CURRENT FILING DATE: 2002-05-07  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 162  
; LENGTH: 343  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-140-807-162

Query Match 40.8%; Score 341; DB 12; Length 343;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGNCGVSIALLVLLGAARLPAGAFAFIAPRESNITVLKGTPTLLAKPCYIVI 60  
Db 1 MAGNCGVSIALLVLLGAARLPAGAFAFIAPRESNITVLKGTPTLLAKPCYIVI 60  
QY 61 SKRHITMISGGERIVFTFSCQSPENHFVIEIQKNIDCMGSPCPGEGVQLQPSSTLLPT 120  
Db 61 SKRHITMISGGERIVFTFSCQSPENHFVIEIQKNIDCMGSPCPGEGVQLQPSSTLLPT 120  
QY 121 LNRFTFMDVKAKHSIGLEQFSIPRLRQIGBESCPDGVTHSIGRIDATVVRIGTFPCSN 180  
Db 121 LNRFTFMDVKAKHSIGLEQFSIPRLRQIGBESCPDGVTHSIGRIDATVVRIGTFPCSN 180  
QY 181 GTVSRIRKQEGVKALHLPWFHPRNVSGFSIANSSIRLCIIISVFEGESATLMSANY 240  
Db 181 GTVSRIRKQEGVKALHLPWFHPRNVSGFSIANSSIRLCIIISVFEGESATLMSANY 240  
QY 241 PEGPPEDELMTWQFVVPALHRAVSFLNFNLSNCRKEERYEYIPGSTTNPVEVKLEDK 300  
Db 241 PEGPPEDELMTWQFVVPALHRAVSFLNFNLSNCRKEERYEYIPGSTTNPVEVKLEDK 300  
QY 301 QPGMAGNFNLSLOGCDODASPGILRLQFOVLVOHPONES 341  
Db 301 QPGMAGNFNLSLOGCDODASPGILRLQFOVLVOHPONES 341

Db 301 QPGMAGNFNLSLOGCDODASPGILRLQFOVLVOHPONES 341  
QY 301 QPGMAGNFNLSLOGCDODASPGILRLQFOVLVOHPONES 341

RESULT 9  
US-10-140-922-162  
; Sequence 162, Application US/10140922  
; Publication No. US20030138889A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: Deforge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Macanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P33081C179  
; CURRENT APPLICATION NUMBER: US/10/140,922  
; CURRENT FILING DATE: 2002-05-07  
; Prior Application removed - See Palm or File Wrapper  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 162  
; LENGTH: 343  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-140-922-162

Query Match 40.8%; Score 341; DB 12; Length 343;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGNCGVSIALLVLLGAARLPAGAFAFIAPRESNITVLKGTPTLLAKPCYIVI 60  
Db 1 MAGNCGVSIALLVLLGAARLPAGAFAFIAPRESNITVLKGTPTLLAKPCYIVI 60  
QY 61 SKRHITMISGGERIVFTFSCQSPENHFVIEIQKNIDCMGSPCPGEGVQLQPSSTLLPT 120  
Db 61 SKRHITMISGGERIVFTFSCQSPENHFVIEIQKNIDCMGSPCPGEGVQLQPSSTLLPT 120  
QY 121 LNRFTFMDVKAKHSIGLEQFSIPRLRQIGBESCPDGVTHSIGRIDATVVRIGTFPCSN 180  
Db 121 LNRFTFMDVKAKHSIGLEQFSIPRLRQIGBESCPDGVTHSIGRIDATVVRIGTFPCSN 180  
QY 181 GTVSRIRKQEGVKALHLPWFHPRNVSGFSIANSSIRLCIIISVFEGESATLMSANY 240  
Db 181 GTVSRIRKQEGVKALHLPWFHPRNVSGFSIANSSIRLCIIISVFEGESATLMSANY 240  
QY 241 PEGPPEDELMTWQFVVPALHRAVSFLNFNLSNCRKEERYEYIPGSTTNPVEVKLEDK 300  
Db 241 PEGPPEDELMTWQFVVPALHRAVSFLNFNLSNCRKEERYEYIPGSTTNPVEVKLEDK 300  
QY 301 QPGMAGNFNLSLOGCDODASPGILRLQFOVLVOHPONES 341  
Db 301 QPGMAGNFNLSLOGCDODASPGILRLQFOVLVOHPONES 341

RESULT 10  
US-10-140-924-162  
; Sequence 162, Application US/10140924  
; Publication No. US20030134355A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.

```

; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey J.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C17
; CURRENT APPLICATION NUMBER: US/10/140,924
; PRIOR APPLICATION REMOVED - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 162
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-924-162

```

```

Query Match      40.8%; Score 341; DB 12; Length 343;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MAGLNCGVSIALLGVLLGAARLPRGAFAFEIALPRESNITVLIKGTPTLLAKCYIYI 60
DB 1 MAGLNCGVSIALLGVLLGAARLPRGAFAFEIALPRESNITVLIKGTPTLLAKCYIYI 60
QY 61 SKRHITMISIKSGERIVTFSCQSPENHFVIEIQKNIDCMGSPCFGEVOLQPSISLPT 120
DB 61 SKRHITMISIKSGERIVTFSCQSPENHFVIEIQKNIDCMGSPCFGEVOLQPSISLPT 120
QY 121 LNRTFIWDVKAHKSIGLEQFSIPRLRQIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
DB 121 LNRTFIWDVKAHKSIGLEQFSIPRLRQIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
QY 181 GTVSRIRKQEGVKALHLPWFHPRVSGFSIANRSSIKRLCTIESEVFEBSGATLMSANY 240
DB 181 GTVSRIRKQEGVKALHLPWFHPRVSGFSIANRSSIKRLCTIESEVFEBSGATLMSANY 240
QY 241 PEGFPEDELMTQFVVPALHRAVSFLLNPNLSNCRKEERYEYIPGSTTNPVEFKLEDK 300
DB 241 PEGFPEDELMTQFVVPALHRAVSFLLNPNLSNCRKEERYEYIPGSTTNPVEFKLEDK 300
QY 301 QPGNMAGNFNLSLQCCDDAOSPGILRLQFVLVQHPONES 341
DB 301 QPGNMAGNFNLSLQCCDDAOSPGILRLQFVLVQHPONES 341

```

# RESULT 11

```

; Sequence 162, Application US/10140926
; Publication No. US20030134356A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey J.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.

```

```

; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C187
; CURRENT APPLICATION NUMBER: US/10/140,926
; PRIOR APPLICATION REMOVED - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 162
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-926-162

```

```

Query Match      40.8%; Score 341; DB 12; Length 343;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MAGLNCGVSIALLGVLLGAARLPRGAFAFEIALPRESNITVLIKGTPTLLAKCYIYI 60
DB 1 MAGLNCGVSIALLGVLLGAARLPRGAFAFEIALPRESNITVLIKGTPTLLAKCYIYI 60
QY 61 SKRHITMISIKSGERIVTFSCQSPENHFVIEIQKNIDCMGSPCFGEVOLQPSISLPT 120
DB 61 SKRHITMISIKSGERIVTFSCQSPENHFVIEIQKNIDCMGSPCFGEVOLQPSISLPT 120
QY 121 LNRTFIWDVKAHKSIGLEQFSIPRLRQIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
DB 121 LNRTFIWDVKAHKSIGLEQFSIPRLRQIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
QY 181 GTVSRIRKQEGVKALHLPWFHPRVSGFSIANRSSIKRLCTIESEVFEBSGATLMSANY 240
DB 181 GTVSRIRKQEGVKALHLPWFHPRVSGFSIANRSSIKRLCTIESEVFEBSGATLMSANY 240
QY 241 PEGFPEDELMTQFVVPALHRAVSFLLNPNLSNCRKEERYEYIPGSTTNPVEFKLEDK 300
DB 241 PEGFPEDELMTQFVVPALHRAVSFLLNPNLSNCRKEERYEYIPGSTTNPVEFKLEDK 300
QY 301 QPGNMAGNFNLSLQCCDDAOSPGILRLQFVLVQHPONES 341
DB 301 QPGNMAGNFNLSLQCCDDAOSPGILRLQFVLVQHPONES 341

```

# RESULT 12

```

; Sequence 162, Application US/10141698
; Publication No. US20030134357A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey J.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: ACIDS ENCODING THE SAME

```

FILE REFERENCE: P3330R1C206  
CURRENT APPLICATION NUMBER: US/10/141.698  
CURRENT FILING DATE: 2002-05-08  
Prior Application removed - See Palm or File Wrapper  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 162  
LENGTH: 343  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-10-141-698-162

Query Match 40.8%; Score 341; DB 12; Length 343;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MAGNCGVSIALGVLGGAARLPRGAFAEIALPRESNITVLKLGPTLLAKRCYIVI 60
DB 1 MAGNCGVSIALGVLGGAARLPRGAFAEIALPRESNITVLKLGPTLLAKRCYIVI 60
QY 61 SKRHITMISIGSERIVFTFCQSPENHFVIEIQKNDMSGPCPFGEVOLQPSISLPT 120
DB 61 SKRHITMISIGSERIVFTFCQSPENHFVIEIQKNDMSGPCPFGEVOLQPSISLPT 120
QY 121 LNRTFIMDVKAHKSIGLEQPSIPRLRQIGPESCPDGVTHSIGRIDATVVRIGTFCSN 180
DB 121 LNRTFIMDVKAHKSIGLEQPSIPRLRQIGPESCPDGVTHSIGRIDATVVRIGTFCSN 180
QY 181 GTVSRIKMQEGVKMALHPWFHPRVNVSFSLNFSNCRKEERYEYIIGSTTNPEVFKLEDK 240
DB 181 GTVSRIKMQEGVKMALHPWFHPRVNVSFSLNFSNCRKEERYEYIIGSTTNPEVFKLEDK 240
QY 241 PEGPPEDELMTQFVVPALHRAVSFLNFNLSNCRKEERYEYIIGSTTNPEVFKLEDK 300
DB 241 PEGPPEDELMTQFVVPALHRAVSFLNFNLSNCRKEERYEYIIGSTTNPEVFKLEDK 300
QY 301 QPGNMGNFNLSLQGCDDQDQSPGILRLQFQVLVQHONES 341
DB 301 QPGNMGNFNLSLQGCDDQDQSPGILRLQFQVLVQHONES 341
```

RESULT 13  
US-10-141-702-162  
Sequence 162, Application US/10141702  
Publication No. US20030134358A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: DeForge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Goddard, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
TITLE OF INVENTION: ACIDS ENCODING THE SAME  
FILE REFERENCE: P3330R1C208  
CURRENT APPLICATION NUMBER: US/10/141.702  
CURRENT FILING DATE: 2002-05-08  
Prior Application removed - See Palm or File Wrapper  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 162  
LENGTH: 343  
TYPE: PRT  
ORGANISM: Homo Sapien

US-10-141-702-162

Query Match 40.8%; Score 341; DB 12; Length 343;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MAGNCGVSIALGVLGGAARLPRGAFAEIALPRESNITVLKLGPTLLAKRCYIVI 60
DB 1 MAGNCGVSIALGVLGGAARLPRGAFAEIALPRESNITVLKLGPTLLAKRCYIVI 60
QY 61 SKRHITMISIGSERIVFTFCQSPENHFVIEIQKNDMSGPCPFGEVOLQPSISLPT 120
DB 61 SKRHITMISIGSERIVFTFCQSPENHFVIEIQKNDMSGPCPFGEVOLQPSISLPT 120
QY 121 LNRTFIMDVKAHKSIGLEQPSIPRLRQIGPESCPDGVTHSIGRIDATVVRIGTFCSN 180
DB 121 LNRTFIMDVKAHKSIGLEQPSIPRLRQIGPESCPDGVTHSIGRIDATVVRIGTFCSN 180
QY 181 GTVSRIKMQEGVKMALHPWFHPRVNVSFSLNFSNCRKEERYEYIIGSTTNPEVFKLEDK 240
DB 181 GTVSRIKMQEGVKMALHPWFHPRVNVSFSLNFSNCRKEERYEYIIGSTTNPEVFKLEDK 240
QY 241 PEGPPEDELMTQFVVPALHRAVSFLNFNLSNCRKEERYEYIIGSTTNPEVFKLEDK 300
DB 241 PEGPPEDELMTQFVVPALHRAVSFLNFNLSNCRKEERYEYIIGSTTNPEVFKLEDK 300
QY 301 QPGNMGNFNLSLQGCDDQDQSPGILRLQFQVLVQHONES 341
DB 301 QPGNMGNFNLSLQGCDDQDQSPGILRLQFQVLVQHONES 341
```

RESULT 14  
US-10-141-704-162  
Sequence 162, Application US/10141704  
Publication No. US20030134359A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: DeForge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Goddard, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
TITLE OF INVENTION: ACIDS ENCODING THE SAME  
FILE REFERENCE: P3330R1C209  
CURRENT APPLICATION NUMBER: US/10/141.704  
CURRENT FILING DATE: 2002-05-08  
Prior Application removed - See Palm or File Wrapper  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 162  
LENGTH: 343  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-10-141-704-162

Query Match 40.8%; Score 341; DB 12; Length 343;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MAGNCGVSIALGVLGGAARLPRGAFAEIALPRESNITVLKLGPTLLAKRCYIVI 60
DB 1 MAGNCGVSIALGVLGGAARLPRGAFAEIALPRESNITVLKLGPTLLAKRCYIVI 60
```

```
QY 61 SRRHITMLSIKSGERIVFTFSCSPENHFVIEIOXNIDCMGSPCPGVEVOLQPSISLPT 120
DB 61 SRRHITMLSIKSGERIVFTFSCSPENHFVIEIOXNIDCMGSPCPGVEVOLQPSISLPT 120
QY 121 LNRFTIWDVKAHKSIGLEIQFSIPRLROIQPGESCPDGVTHSISGRIDATVVRIGTFCGN 180
DB 121 LNRFTIWDVKAHKSIGLEIQFSIPRLROIQPGESCPDGVTHSISGRIDATVVRIGTFCGN 180
QY 181 GTVSRIKMQEGVKMALHLPWFHPRNVSGFSIANRSSIKRLCTIESVFEGESATLMSANY 240
DB 181 GTVSRIKMQEGVKMALHLPWFHPRNVSGFSIANRSSIKRLCTIESVFEGESATLMSANY 240
QY 241 PEGFPEDELMTQFVVPALHRAVSFLNPNLSNCRKEERYEYYPGSTTNPVEVKLEBK 300
DB 241 PEGFPEDELMTQFVVPALHRAVSFLNPNLSNCRKEERYEYYPGSTTNPVEVKLEBK 300
QY 301 QPGNMAGNPNLSLOGCDDAQSPGILRLQFVLVQHPONES 341
DB 301 QPGNMAGNPNLSLOGCDDAQSPGILRLQFVLVQHPONES 341
```

## RESULT 15

```
US-10-142-421-162
; Sequence 162, Application US/10142421
; Publication No. US20030134360A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvarolf, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Guiney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C218
; CURRENT APPLICATION NUMBER: US/10/142,421
; CURRENT FILING DATE: 2002-05-09
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 162
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-142-421-162
```

```
Query Match 40.8%; Score 341; DB 12; Length 343;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MAGLNCGVSIALLGVLLGAARLPKGAFAEIALPRESNITVLIKIGTPTLLAKPCYIV 60
DB 1 MAGLNCGVSIALLGVLLGAARLPKGAFAEIALPRESNITVLIKIGTPTLLAKPCYIV 60
QY 61 SRRHITMLSIKSGERIVFTFSCSPENHFVIEIOXNIDCMGSPCPGVEVOLQPSISLPT 120
DB 61 SRRHITMLSIKSGERIVFTFSCSPENHFVIEIOXNIDCMGSPCPGVEVOLQPSISLPT 120
QY 121 LNRFTIWDVKAHKSIGLEIQFSIPRLROIQPGESCPDGVTHSISGRIDATVVRIGTFCGN 180
DB 121 LNRFTIWDVKAHKSIGLEIQFSIPRLROIQPGESCPDGVTHSISGRIDATVVRIGTFCGN 180
```

```
QY 181 GTVSRIKMQEGVKMALHLPWFHPRNVSGFSIANRSSIKRLCTIESVFEGESATLMSANY 240
DB 181 GTVSRIKMQEGVKMALHLPWFHPRNVSGFSIANRSSIKRLCTIESVFEGESATLMSANY 240
QY 241 PEGFPEDELMTQFVVPALHRAVSFLNPNLSNCRKEERYEYYPGSTTNPVEVKLEBK 300
DB 241 PEGFPEDELMTQFVVPALHRAVSFLNPNLSNCRKEERYEYYPGSTTNPVEVKLEBK 300
QY 301 QPGNMAGNPNLSLOGCDDAQSPGILRLQFVLVQHPONES 341
DB 301 QPGNMAGNPNLSLOGCDDAQSPGILRLQFVLVQHPONES 341
```

Search completed: February 20, 2004, 19:52:51  
Job time : 102 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using -bw model

Run on: February 20, 2004, 19:47:18 ; Search time 360 Seconds

(without alignments)  
2113.034 Million cell updates/sec

US-09-899-569A-4

Title: 836  
Perfect score: 1 MAGNCGSIALLCVLLGA.....SSKDTDIPLANTORPMEPAE 836  
Sequence:

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 5728757 seqs, 909918778 residues

Word size: 0

Total number of hits satisfying chosen parameters: 5728757

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

Pending\_Patents\_AA\_Main:\*

1: /cgn2\_6/ptodata/1/paa/PCRTUS\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/paa/US06\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/paa/US07\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/paa/US081\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/paa/US082\_COMB.pep:\*  
7: /cgn2\_6/ptodata/1/paa/US083\_COMB.pep:\*  
8: /cgn2\_6/ptodata/1/paa/US084\_COMB.pep:\*  
9: /cgn2\_6/ptodata/1/paa/US085\_COMB.pep:\*  
10: /cgn2\_6/ptodata/1/paa/US086\_COMB.pep:\*  
11: /cgn2\_6/ptodata/1/paa/US087\_COMB.pep:\*  
12: /cgn2\_6/ptodata/1/paa/US088\_COMB.pep:\*  
13: /cgn2\_6/ptodata/1/paa/US089\_COMB.pep:\*  
14: /cgn2\_6/ptodata/1/paa/US090\_COMB.pep:\*  
15: /cgn2\_6/ptodata/1/paa/US091\_COMB.pep:\*  
16: /cgn2\_6/ptodata/1/paa/US092\_COMB.pep:\*  
17: /cgn2\_6/ptodata/1/paa/US093\_COMB.pep:\*  
18: /cgn2\_6/ptodata/1/paa/US094\_COMB.pep:\*  
19: /cgn2\_6/ptodata/1/paa/US095\_COMB.pep:\*  
20: /cgn2\_6/ptodata/1/paa/US096\_COMB.pep:\*  
21: /cgn2\_6/ptodata/1/paa/US097\_COMB.pep:\*  
22: /cgn2\_6/ptodata/1/paa/US097B\_COMB.pep:\*  
23: /cgn2\_6/ptodata/1/paa/US098\_COMB.pep:\*  
24: /cgn2\_6/ptodata/1/paa/US099\_COMB.pep:\*  
25: /cgn2\_6/ptodata/1/paa/US099B\_COMB.pep:\*  
26: /cgn2\_6/ptodata/1/paa/US099C\_COMB.pep:\*  
27: /cgn2\_6/ptodata/1/paa/US100\_COMB.pep:\*  
28: /cgn2\_6/ptodata/1/paa/US101\_COMB.pep:\*  
29: /cgn2\_6/ptodata/1/paa/US102\_COMB.pep:\*  
30: /cgn2\_6/ptodata/1/paa/US103\_COMB.pep:\*  
31: /cgn2\_6/ptodata/1/paa/US104\_COMB.pep:\*  
32: /cgn2\_6/ptodata/1/paa/US106\_COMB.pep:\*  
33: /cgn2\_6/ptodata/1/paa/US106\_COMB.pep:\*  
34: /cgn2\_6/ptodata/1/paa/US106\_COMB.pep:\*  
35: /cgn2\_6/ptodata/1/paa/US106\_COMB.pep:\*  
36: /cgn2\_6/ptodata/1/paa/US106\_COMB.pep:\*  
37: /cgn2\_6/ptodata/1/paa/US106\_COMB.pep:\*  
38: /cgn2\_6/ptodata/1/paa/US106\_COMB.pep:\*  
39: /cgn2\_6/ptodata/1/paa/US106\_COMB.pep:\*  
40: /cgn2\_6/ptodata/1/paa/US106\_COMB.pep:\*  
41: /cgn2\_6/ptodata/1/paa/US106\_COMB.pep:\*  
42: /cgn2\_6/ptodata/1/paa/US106\_COMB.pep:\*  
43: /cgn2\_6/ptodata/1/paa/US106\_COMB.pep:\*  
44: /cgn2\_6/ptodata/1/paa/US106\_COMB.pep:\*  
45: /cgn2\_6/ptodata/1/paa/US106\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	836	100.0	836 1	PCT-US02-05095A-1600 Sequence 1600, App

2	836	100.0	836	23	US-09-899-569A-4	Sequence 4, Appl
3	650	77.8	749	23	US-09-899-569A-2	Sequence 2, Appl
4	413	49.4	414	29	US-10-351-334-273	Sequence 273, App
5	413	49.4	443	1	PCT-US99-17130-265	Sequence 265, App
6	413	49.4	443	29	US-10-351-334-271	Sequence 271, App
7	382	45.7	649	20	US-09-629-469A-18191	Sequence 18191, A
8	341	40.8	343	1	PCT-US01-43523-162	Sequence 162, App
9	341	40.8	343	1	PCT-US02-24563-162	Sequence 162, App
10	341	40.8	343	26	US-10-028-072-162	Sequence 162, App
11	341	40.8	343	27	US-10-121-040-162	Sequence 162, App
12	341	40.8	343	27	US-10-121-041-162	Sequence 162, App
13	341	40.8	343	27	US-10-121-042-162	Sequence 162, App
14	341	40.8	343	27	US-10-121-043-162	Sequence 162, App
15	341	40.8	343	27	US-10-121-044-162	Sequence 162, App
16	341	40.8	343	27	US-10-121-045-162	Sequence 162, App
17	341	40.8	343	27	US-10-121-046-162	Sequence 162, App
18	341	40.8	343	27	US-10-121-047-162	Sequence 162, App
19	341	40.8	343	27	US-10-121-048-162	Sequence 162, App
20	341	40.8	343	27	US-10-121-049-162	Sequence 162, App
21	341	40.8	343	27	US-10-121-050-162	Sequence 162, App
22	341	40.8	343	27	US-10-121-051-162	Sequence 162, App
23	341	40.8	343	27	US-10-121-052-162	Sequence 162, App
24	341	40.8	343	27	US-10-121-053-162	Sequence 162, App
25	341	40.8	343	27	US-10-121-054-162	Sequence 162, App
26	341	40.8	343	27	US-10-121-055-162	Sequence 162, App
27	341	40.8	343	27	US-10-121-056-162	Sequence 162, App
28	341	40.8	343	27	US-10-121-057-162	Sequence 162, App
29	341	40.8	343	27	US-10-121-058-162	Sequence 162, App
30	341	40.8	343	27	US-10-121-059-162	Sequence 162, App
31	341	40.8	343	27	US-10-121-060-162	Sequence 162, App
32	341	40.8	343	27	US-10-121-061-162	Sequence 162, App
33	341	40.8	343	27	US-10-121-062-162	Sequence 162, App
34	341	40.8	343	27	US-10-123-108-162	Sequence 162, App
35	341	40.8	343	27	US-10-123-109-162	Sequence 162, App
36	341	40.8	343	27	US-10-123-154-162	Sequence 162, App
37	341	40.8	343	27	US-10-123-156-162	Sequence 162, App
38	341	40.8	343	27	US-10-123-157-162	Sequence 162, App
39	341	40.8	343	27	US-10-123-212-162	Sequence 162, App
40	341	40.8	343	27	US-10-123-213-162	Sequence 162, App
41	341	40.8	343	27	US-10-123-214-162	Sequence 162, App
42	341	40.8	343	27	US-10-123-215-162	Sequence 162, App
43	341	40.8	343	27	US-10-123-235-162	Sequence 162, App
44	341	40.8	343	27	US-10-123-236-162	Sequence 162, App
45	341	40.8	343	27	US-10-123-261-162	Sequence 162, App

ALIGNMENTS

RESULT 1  
PCT-US02-05095A-1600  
Sequence 1600, Application PC/TUS0205095A  
GENERAL INFORMATION:  
APPLICANT: Hyseq, Inc  
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
FILE REFERENCE: 21272-075 (803)  
CURRENT APPLICATION NUMBER: PCT/US02/05095A  
PRIOR FILING DATE: 2002-03-05  
PRIOR APPLICATION NUMBER: 09/799,451  
NUMBER OF SEQ ID NOS: 1896  
SEQ ID NO 1600  
LENGTH: 836  
TYPE: PRT  
ORGANISM: Homo sapiens  
PCT-US02-05095A-1600

Query Match 100.0%; Score 836; DB 1; Length 836;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 836; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 MAGNCGSIALLCVLLGAARLPAGAAFEIALPRESNTIVLTKLGPTTLARPCVVI 60  
|||||

```
Db 1 MAGNCGVSIALGVLLGAARLPGRGAFAETALPRESNITVLIKLGTPTLLAKPCYIV 60
Qy 61 SKRHITMLSIKSGERIVTFPSCQSPNNHFVIEIQKNDICMSGPCPGEVOLQPSISLPT 120
Db 61 SKRHITMLSIKSGERIVTFPSCQSPNNHFVIEIQKNDICMSGPCPGEVOLQPSISLPT 120
Qy 121 LNRFTFMDVKAHKSIGLELOFSIPRLROIIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
Db 121 LNRFTFMDVKAHKSIGLELOFSIPRLROIIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
Qy 181 GTVSRIKMOEGVKMALHLPWFHPRNVSGFSIANRSSIKRLCIIESVFESEGSATLSMANY 240
Db 181 GTVSRIKMOEGVKMALHLPWFHPRNVSGFSIANRSSIKRLCIIESVFESEGSATLSMANY 240
Qy 241 PEGFPEDELMQOFVVPALHRAVSFLNPNLSNCKEKEVEYYIPGSTTNEVEFKLEDK 300
Db 241 PEGFPEDELMQOFVVPALHRAVSFLNPNLSNCKEKEVEYYIPGSTTNEVEFKLEDK 300
Qy 301 QGNNAGNFNLSLOGCDQDQASPGILRLQFQVLAQHPQESNKIYVVDLSNERAMSLTIE 360
Db 301 QGNNAGNFNLSLOGCDQDQASPGILRLQFQVLAQHPQESNKIYVVDLSNERAMSLTIE 360
Qy 361 PRPVQSRKFVPGCFVCESTRCSNLTLSGSKHKISFLCDDLRLMANNVEKTISCTDH 420
Db 361 PRPVQSRKFVPGCFVCESTRCSNLTLSGSKHKISFLCDDLRLMANNVEKTISCTDH 420
Qy 421 RYCQKRSYSLOVPSDILHLPELHDFSWKLLVPKDRSLVLPVPAQKLOQTHEKPCNTSF 480
Db 421 RYCQKRSYSLOVPSDILHLPELHDFSWKLLVPKDRSLVLPVPAQKLOQTHEKPCNTSF 480
Qy 481 SYLVASAIPSODLTFGSCPCGSGIKQIQVKONISVTLRTFAPSFOEASRQGLTVSFTFY 540
Db 481 SYLVASAIPSODLTFGSCPCGSGIKQIQVKONISVTLRTFAPSFOEASRQGLTVSFTFY 540
Qy 541 FKEBGFVTPPTKSKVYLRTPNMDRGLPSLTSVSWNISVPRDOVACLTFEKERSGVCC 600
Db 541 FKEBGFVTPPTKSKVYLRTPNMDRGLPSLTSVSWNISVPRDOVACLTFEKERSGVCC 600
Qy 601 TGRAFMIIQEORTAEIIFSLDEEDVLPKPSFHHHSFWNINSCSPTSKQDLFLFSVTLT 660
Db 601 TGRAFMIIQEORTAEIIFSLDEEDVLPKPSFHHHSFWNINSCSPTSKQDLFLFSVTLT 660
Qy 661 PRVVDLTVLLAAVGGVLLSALGLIICVKKKKKTKNKPAGVINYGNINTEMPROPK 720
Db 661 PRVVDLTVLLAAVGGVLLSALGLIICVKKKKKTKNKPAGVINYGNINTEMPROPK 720
Qy 721 KFOKGRKNDSHVAVIEDTWYGHLLQDSGSLQPEVDYTRPQGTWGCPSPTTIC 780
Db 721 KFOKGRKNDSHVAVIEDTWYGHLLQDSGSLQPEVDYTRPQGTWGCPSPTTIC 780
Qy 781 SRAPTAKLATEEPPRSPSESESEPTTFSHPNNGDVSSKOTDIPILNTQOEPMEPAE 836
Db 781 SRAPTAKLATEEPPRSPSESESEPTTFSHPNNGDVSSKOTDIPILNTQOEPMEPAE 836

RESULT 2
US-09-899-569A-4
; Sequence 4, Application US/09899569A
; GENERAL INFORMATION:
; APPLICANT: Norbert Schweifer
; APPLICANT: Marwa Scherl-Mostageer
; APPLICANT: Wolfgang Sommergruber
; APPLICANT: Roger Abseher
; TITLE OF INVENTION: Tumorsassoziierter Antigen (B345)
; FILE REFERENCE: 0652.2280001
; CURRENT APPLICATION NUMBER: US/09/899,569A
; CURRENT FILING DATE: 2001-07-06
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: DE 100 33 080.0
; PRIOR FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: US 60/243,158
; PRIOR FILING DATE: 2000-10-25
```

```
; PRIOR APPLICATION NUMBER: US 60/297,747
; PRIOR FILING DATE: 2001-06-14
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 836
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-899-569A-4

Query Match 100.0%; Score 836; DB 23; Length 836;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 836; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAGNCGVSIALGVLLGAARLPGRGAFAETALPRESNITVLIKLGTPTLLAKPCYIV 60
Db 1 MAGNCGVSIALGVLLGAARLPGRGAFAETALPRESNITVLIKLGTPTLLAKPCYIV 60
Qy 61 SKRHITMLSIKSGERIVTFPSCQSPNNHFVIEIQKNDICMSGPCPGEVOLQPSISLPT 120
Db 61 SKRHITMLSIKSGERIVTFPSCQSPNNHFVIEIQKNDICMSGPCPGEVOLQPSISLPT 120
Qy 121 LNRFTFMDVKAHKSIGLELOFSIPRLROIIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
Db 121 LNRFTFMDVKAHKSIGLELOFSIPRLROIIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
Qy 181 GTVSRIKMOEGVKMALHLPWFHPRNVSGFSIANRSSIKRLCIIESVFESEGSATLSMANY 240
Db 181 GTVSRIKMOEGVKMALHLPWFHPRNVSGFSIANRSSIKRLCIIESVFESEGSATLSMANY 240
Qy 241 PEGFPEDELMQOFVVPALHRAVSFLNPNLSNCKEKEVEYYIPGSTTNEVEFKLEDK 300
Db 241 PEGFPEDELMQOFVVPALHRAVSFLNPNLSNCKEKEVEYYIPGSTTNEVEFKLEDK 300
Qy 301 QGNNAGNFNLSLOGCDQDQASPGILRLQFQVLAQHPQESNKIYVVDLSNERAMSLTIE 360
Db 301 QGNNAGNFNLSLOGCDQDQASPGILRLQFQVLAQHPQESNKIYVVDLSNERAMSLTIE 360
Qy 361 PRPVQSRKFVPGCFVCESTRCSNLTLSGSKHKISFLCDDLRLMANNVEKTISCTDH 420
Db 361 PRPVQSRKFVPGCFVCESTRCSNLTLSGSKHKISFLCDDLRLMANNVEKTISCTDH 420
Qy 421 RYCQKRSYSLOVPSDILHLPELHDFSWKLLVPKDRSLVLPVPAQKLOQTHEKPCNTSF 480
Db 421 RYCQKRSYSLOVPSDILHLPELHDFSWKLLVPKDRSLVLPVPAQKLOQTHEKPCNTSF 480
Qy 481 SYLVASAIPSODLTFGSCPCGSGIKQIQVKONISVTLRTFAPSFOEASRQGLTVSFTFY 540
Db 481 SYLVASAIPSODLTFGSCPCGSGIKQIQVKONISVTLRTFAPSFOEASRQGLTVSFTFY 540
Qy 541 FKEBGFVTPPTKSKVYLRTPNMDRGLPSLTSVSWNISVPRDOVACLTFEKERSGVCC 600
Db 541 FKEBGFVTPPTKSKVYLRTPNMDRGLPSLTSVSWNISVPRDOVACLTFEKERSGVCC 600
Qy 601 TGRAFMIIQEORTAEIIFSLDEEDVLPKPSFHHHSFWNINSCSPTSKQDLFLFSVTLT 660
Db 601 TGRAFMIIQEORTAEIIFSLDEEDVLPKPSFHHHSFWNINSCSPTSKQDLFLFSVTLT 660
Qy 661 PRVVDLTVLLAAVGGVLLSALGLIICVKKKKKTKNKPAGVINYGNINTEMPROPK 720
Db 661 PRVVDLTVLLAAVGGVLLSALGLIICVKKKKKTKNKPAGVINYGNINTEMPROPK 720
Qy 721 KFOKGRKNDSHVAVIEDTWYGHLLQDSGSLQPEVDYTRPQGTWGCPSPTTIC 780
Db 721 KFOKGRKNDSHVAVIEDTWYGHLLQDSGSLQPEVDYTRPQGTWGCPSPTTIC 780
Qy 781 SRAPTAKLATEEPPRSPSESESEPTTFSHPNNGDVSSKOTDIPILNTQOEPMEPAE 836
Db 781 SRAPTAKLATEEPPRSPSESESEPTTFSHPNNGDVSSKOTDIPILNTQOEPMEPAE 836

RESULT 3
US-09-899-569A-2
```



```
; Sequence 2, Application US/09899569A
; GENERAL INFORMATION:
; APPLICANT: Norbert Schweifer
; APPLICANT: Marwa Scherl-Mostaguer
; APPLICANT: Wolfgang Sommergruber
; APPLICANT: Roger Abbeher
; TITLE OF INVENTION: Tumorsoziiertes Antigen (B345)
; FILE REFERENCE: 0652.2280001
; CURRENT APPLICATION NUMBER: US/09/899,569A
; PRIOR FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: DE 100 33 080.0
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: DE 101 19 294.0
; PRIOR FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: US 60/243,158
; PRIOR FILING DATE: 2000-10-25
; PRIOR APPLICATION NUMBER: US 60/297,747
; PRIOR FILING DATE: 2001-06-14
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 749
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-899-569A-2

Query Match      77.8%; Score 650; DB 23; Length 749;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 650; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 MLSTKGRIVTFSCQSPENHFVIEIQKIDCMGSPCPFGVQLQPSSTLLPTLNRTFI 126
DB 1 MLSTKGRIVTFSCQSPENHFVIEIQKIDCMGSPCPFGVQLQPSSTLLPTLNRTFI 60
QY 127 MDVAKHKSIGLELOPSIRLRQIGGSCPDGVTHTSIGRIDATVVRIGTFCSCNTVSRI 186
DB 61 MDVAKHKSIGLELOPSIRLRQIGGSCPDGVTHTSIGRIDATVVRIGTFCSCNTVSRI 120
QY 187 KMOEGVKALHLPMFHPNVSQFSIANKSSIRKLCIISVFEGBSATLMSANYPEGPE 246
DB 121 KMOEGVKALHLPMFHPNVSQFSIANKSSIRKLCIISVFEGBSATLMSANYPEGPE 180
QY 247 DELMTQGVNPAHLASVSLFNFLNSCERKEERYIIPGTTNPEVFKLEDKOPGMA 306
DB 181 DELMTQGVNPAHLASVSLFNFLNSCERKEERYIIPGTTNPEVFKLEDKOPGMA 240
QY 307 GNFNLSTGCCDDAOSPGLRLQFOVLVQHPONESNKIYVVDLSNERAMSLTIEPRVKQ 366
DB 241 GNFNLSTGCCDDAOSPGLRLQFOVLVQHPONESNKIYVVDLSNERAMSLTIEPRVKQ 300
QY 367 SRKPVPGCFVCLSESTCSSNLTLTSGSKHKSIFLCDDLTRLMNVEXTISCDDHRYCQRK 426
DB 301 SRKPVPGCFVCLSESTCSSNLTLTSGSKHKSIFLCDDLTRLMNVEXTISCDDHRYCQRK 360
QY 427 SYSLOVPSDILHLPELHDFSWKLLVPRKRLSLVLVPAKLOQHTHEKPCNTSFSYLVA 486
DB 361 SYSLOVPSDILHLPELHDFSWKLLVPRKRLSLVLVPAKLOQHTHEKPCNTSFSYLVA 420
QY 487 AIPSDOLYFGSFCPEGSIKQIOVKONISVTLTFAPSFOEASRQGLTVSFIYKKEGV 546
DB 421 AIPSDOLYFGSFCPEGSIKQIOVKONISVTLTFAPSFOEASRQGLTVSFIYKKEGV 480
QY 547 FTVPDTSKYVLRTPMNDRGLPSLTYSWINSVPRDVACTPFKESGVAVCCQGRAFM 606
DB 481 FTVPDTSKYVLRTPMNDRGLPSLTYSWINSVPRDVACTPFKESGVAVCCQGRAFM 540
QY 607 IIOEORTAEERISLDEVLPRKPSFHHSFWNINSNCSPTSGKODLLFSVTLTPTVDL 666
DB 541 IIOEORTAEERISLDEVLPRKPSFHHSFWNINSNCSPTSGKODLLFSVTLTPTVDL 600
QY 667 TVILIAAVGGVLLLSALGLIICVKKKKKTKNGPAVGIVNGINTEMP 716
DB 601 TVILIAAVGGVLLLSALGLIICVKKKKKTKNGPAVGIVNGINTEMP 650
```

```
RESULT 4
US-10-351-334-273
; Sequence 273, Application US/10351334
; GENERAL INFORMATION:
; APPLICANT: Komatsoulis et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031P2
; CURRENT APPLICATION NUMBER: US/10/351,334
; PRIOR FILING DATE: 2003-01-27
; PRIOR APPLICATION NUMBER: 60/350,898
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/489,847
; PRIOR FILING DATE: 2000-01-24
; PRIOR APPLICATION NUMBER: PCT/US99/17130
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: 60/094,657
; PRIOR FILING DATE: 1998-07-30
; PRIOR APPLICATION NUMBER: 60/095,486
; PRIOR FILING DATE: 1998-08-05
; PRIOR APPLICATION NUMBER: 60/096,319
; PRIOR FILING DATE: 1998-08-12
; PRIOR APPLICATION NUMBER: 60/095,454
; PRIOR FILING DATE: 1998-08-06
; PRIOR APPLICATION NUMBER: 60/095,455
; PRIOR FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 273
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-351-334-273

Query Match      49.4%; Score 413; DB 29; Length 414;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGINCQVSIALLGVLLGAARLRPGAERIALPRESENTIVLKLGTPTLLARCYVI 60
DB 1 MAGINCQVSIALLGVLLGAARLRPGAERIALPRESENTIVLKLGTPTLLARCYVI 60
QY 61 SKRHITMLSTKGRIVTFSCQSPENHFVIEIQKIDCMGSPCPFGVQLQPSSTLLPT 120
DB 61 SKRHITMLSTKGRIVTFSCQSPENHFVIEIQKIDCMGSPCPFGVQLQPSSTLLPT 120
QY 121 LNRFTIMDVKAHKSIGLELOPSIRLRQIGGSCPDGVTHTSIGRIDATVVRIGTFCSCN 180
DB 121 LNRFTIMDVKAHKSIGLELOPSIRLRQIGGSCPDGVTHTSIGRIDATVVRIGTFCSCN 180
QY 181 GTVSRIRKMOEGVKALHLPMFHPNVSQFSIANKSSIRKLCIISVFEGBSATLMSANY 240
DB 181 GTVSRIRKMOEGVKALHLPMFHPNVSQFSIANKSSIRKLCIISVFEGBSATLMSANY 240
QY 241 PEGFPDELMTQGVNPAHLASVSLFNFLNSCERKEERYIIPGTTNPEVFKLEDK 300
DB 241 PEGFPDELMTQGVNPAHLASVSLFNFLNSCERKEERYIIPGTTNPEVFKLEDK 300
QY 301 QPGMAGNFNLSTGCCDDAOSPGLRLQFOVLVQHPONESNKIYVVDLSNERAMSLTIE 360
DB 301 QPGMAGNFNLSTGCCDDAOSPGLRLQFOVLVQHPONESNKIYVVDLSNERAMSLTIE 360
QY 361 PRPVQSRKPVPGCFVCLSESTCSSNLTLTSGSKHKSIFLCDDLTRLMNVYK 413
DB 361 PRPVQSRKPVPGCFVCLSESTCSSNLTLTSGSKHKSIFLCDDLTRLMNVYK 413

RESULT 5
PCT-US99-17130-265
; Sequence 265, Application PC/TUS9917130
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
```

;; TITLE OF INVENTION: 98 Human Secreted Proteins  
;; FILE REFERENCE: P2031.PCT  
;; CURRENT APPLICATION NUMBER: PCT/US99/17130  
;; CURRENT FILING DATE: 1999-07-29  
;; EARLIER APPLICATION NUMBER: 60/094,457  
;; EARLIER FILING DATE: 1998-07-30  
;; EARLIER APPLICATION NUMBER: 60/095,486  
;; EARLIER FILING DATE: 1998-08-05  
;; EARLIER APPLICATION NUMBER: 60/095,455  
;; EARLIER FILING DATE: 1998-08-06  
;; EARLIER APPLICATION NUMBER: 60/095,454  
;; EARLIER FILING DATE: 1998-08-06  
;; EARLIER APPLICATION NUMBER: 60/096,319  
;; EARLIER FILING DATE: 1998-08-12  
;; NUMBER OF SEQ ID NOS: 364  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 265  
;; LENGTH: 443  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
PCT-US99-17130-265

Query Match 49.4%; Score 413; DB 1; Length 443;

Best Local Similarity 100.0%; Pred. No. 0; Matches 413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGLNCVSIALGLVLLGAARLPRGAFAFEIALPRESNITVLIKGTPTLLAKPCYIV 60  
DB 30 MAGLNCVSIALGLVLLGAARLPRGAFAFEIALPRESNITVLIKGTPTLLAKPCYIV 89  
QY 61 SKRHTMTLSKSGEIVTFPSCSPENHFVIEIQKNIDMSGCPGGEVQLOPSTSLPT 120  
DB 90 SKRHTMTLSKSGEIVTFPSCSPENHFVIEIQKNIDMSGCPGGEVQLOPSTSLPT 149  
QY 121 LNRTFIMVKAHKSIGLELOFSIPRLROIQPGSCPDGVTSHISGRIDATVIRIGTFCSN 180  
DB 150 LNRTFIMVKAHKSIGLELOFSIPRLROIQPGSCPDGVTSHISGRIDATVIRIGTFCSN 209  
QY 181 GTVSRIKMOEGVKMALHPMFHPRNVSGFSIANRSSIKRLCTIESEVFEGEGSATLMSANY 240  
DB 210 GTVSRIKMOEGVKMALHPMFHPRNVSGFSIANRSSIKRLCTIESEVFEGEGSATLMSANY 269  
QY 241 PGGFPEDELTMTQFVVPALRLASVSFLNPNLSNCKEKEKREVEYYIPGSTTNEVEFKLEBK 300  
DB 270 PGGFPEDELTMTQFVVPALRLASVSFLNPNLSNCKEKEKREVEYYIPGSTTNEVEFKLEBK 329  
QY 301 QPGNAGNFNLISLQCGDDAOSPGLRLQFOVLVQHPONESNKIYVVDLSNERAMSLTTE 360  
DB 330 QPGNAGNFNLISLQCGDDAOSPGLRLQFOVLVQHPONESNKIYVVDLSNERAMSLTTE 389  
QY 361 PRPVKSRKRVPGCFVCLSESRTCSSNLTLTSGSKHISFLCDLTRLMMNVK 413  
DB 390 PRPVKSRKRVPGCFVCLSESRTCSSNLTLTSGSKHISFLCDLTRLMMNVK 442

RESULT 6  
US-10-351-334-271  
; Sequence 271, Application US/10351334  
; GENERAL INFORMATION:

;; APPLICANT: Komatsushita et al  
;; TITLE OF INVENTION: 98 Human Secreted Proteins  
;; FILE REFERENCE: P2031P2  
;; CURRENT APPLICATION NUMBER: US/10/351,334  
;; CURRENT FILING DATE: 2003-01-27  
;; PRIOR APPLICATION NUMBER: 60/350,898  
;; PRIOR FILING DATE: 2002-01-25  
;; PRIOR APPLICATION NUMBER: 09/489,847  
;; PRIOR FILING DATE: 2000-01-24  
;; PRIOR APPLICATION NUMBER: PCT/US99/17130  
;; PRIOR FILING DATE: 1999-07-29  
;; PRIOR APPLICATION NUMBER: 60/094,657  
;; PRIOR FILING DATE: 1998-07-30  
;; PRIOR APPLICATION NUMBER: 60/095,486

;; PRIOR FILING DATE: 1998-08-05  
;; PRIOR APPLICATION NUMBER: 60/096,319  
;; PRIOR FILING DATE: 1998-08-12  
;; PRIOR APPLICATION NUMBER: 60/095,454  
;; PRIOR FILING DATE: 1998-08-06  
;; PRIOR APPLICATION NUMBER: 60/095,455  
;; PRIOR FILING DATE: 1998-08-06  
;; NUMBER OF SEQ ID NOS: 376  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 271  
;; LENGTH: 443  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-351-334-271

Query Match 49.4%; Score 413; DB 29; Length 443;

Best Local Similarity 100.0%; Pred. No. 0; Matches 413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGLNCVSIALGLVLLGAARLPRGAFAFEIALPRESNITVLIKGTPTLLAKPCYIV 60  
DB 30 MAGLNCVSIALGLVLLGAARLPRGAFAFEIALPRESNITVLIKGTPTLLAKPCYIV 89  
QY 61 SKRHTMTLSKSGEIVTFPSCSPENHFVIEIQKNIDMSGCPGGEVQLOPSTSLPT 120  
DB 90 SKRHTMTLSKSGEIVTFPSCSPENHFVIEIQKNIDMSGCPGGEVQLOPSTSLPT 149  
QY 121 LNRTFIMVKAHKSIGLELOFSIPRLROIQPGSCPDGVTSHISGRIDATVIRIGTFCSN 180  
DB 150 LNRTFIMVKAHKSIGLELOFSIPRLROIQPGSCPDGVTSHISGRIDATVIRIGTFCSN 209  
QY 181 GTVSRIKMOEGVKMALHPMFHPRNVSGFSIANRSSIKRLCTIESEVFEGEGSATLMSANY 240  
DB 210 GTVSRIKMOEGVKMALHPMFHPRNVSGFSIANRSSIKRLCTIESEVFEGEGSATLMSANY 269  
QY 241 PGGFPEDELTMTQFVVPALRLASVSFLNPNLSNCKEKEKREVEYYIPGSTTNEVEFKLEBK 300  
DB 270 PGGFPEDELTMTQFVVPALRLASVSFLNPNLSNCKEKEKREVEYYIPGSTTNEVEFKLEBK 329  
QY 301 QPGNAGNFNLISLQCGDDAOSPGLRLQFOVLVQHPONESNKIYVVDLSNERAMSLTTE 360  
DB 330 QPGNAGNFNLISLQCGDDAOSPGLRLQFOVLVQHPONESNKIYVVDLSNERAMSLTTE 389  
QY 361 PRPVKSRKRVPGCFVCLSESRTCSSNLTLTSGSKHISFLCDLTRLMMNVK 413  
DB 390 PRPVKSRKRVPGCFVCLSESRTCSSNLTLTSGSKHISFLCDLTRLMMNVK 442

RESULT 7  
US-09-629-469A-18191  
; Sequence 18191, Application US/09629469A  
; GENERAL INFORMATION:

;; APPLICANT: OITA, TOSHIO  
;; APPLICANT: ISOGAI, TAKAO  
;; APPLICANT: NISHIKAWA, TETSUO  
;; APPLICANT: HAYASHI, KOJI  
;; APPLICANT: SAITO, KAORU  
;; APPLICANT: YAMAMOTO, JUNICHI  
;; APPLICANT: ISHII, SHIZUKO  
;; APPLICANT: SUGIYAMA, TOMOYASU  
;; APPLICANT: WAKAMATSU, AI  
;; APPLICANT: NAGAI, KEIICHI  
;; APPLICANT: OTSUKI, TETSUJI  
;; TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH cDNA AND THEIR USE  
;; FILE REFERENCE: 084335/0123  
;; CURRENT APPLICATION NUMBER: US/09/629,469A  
;; CURRENT FILING DATE: 2000-07-28  
;; PRIOR APPLICATION NUMBER: JP 1999-248036  
;; PRIOR FILING DATE: 1999-07-29  
;; PRIOR APPLICATION NUMBER: JP 1999-300253  
;; PRIOR FILING DATE: 1999-08-27  
;; PRIOR APPLICATION NUMBER: JP 2000-118776  
;; PRIOR FILING DATE: 2000-01-11

PRIOR APPLICATION NUMBER: JP 2000-183767  
PRIOR FILING DATE: 2000-05-02  
PRIOR APPLICATION NUMBER: JP 2000-241899  
PRIOR FILING DATE: 2000-06-09  
PRIOR APPLICATION NUMBER: 60/159,590  
PRIOR FILING DATE: 1999-10-18  
PRIOR APPLICATION NUMBER: 60/183,322  
PRIOR FILING DATE: 2000-02-17  
NUMBER OF SEQ ID NOS: 19025  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO: 18191  
LENGTH: 649  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-629-469A-18191

Query Match 45.7%; Score 382; DB 20; Length 649;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 582; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

253 QFVPAHLRASVFLNPNLNCERKEERYIIPGTTPEVFKLEDKOPGMAGNPNLS 312  
66 QFVPAHLRASVFLNPNLNCERKEERYIIPGTTPEVFKLEDKOPGMAGNPNLS 125  
313 LOGCDQDASPGILRLQFVLVQHPONESNKIYVVDLSNERAMSLTEBRPVKQSRKFPV 372  
126 LOGCDQDASPGILRLQFVLVQHPONESNKIYVVDLSNERAMSLTEBRPVKQSRKFPV 185  
373 GGFVLESTGSSNLTLSGSGHKSIFLCCDLTILMANVEKTICTDHRVCKRKSISLOY 432  
186 GGFVLESTGSSNLTLSGSGHKSIFLCCDLTILMANVEKTICTDHRVCKRKSISLOY 245  
433 PEDILHLPEVLEHDFGSKLIVPKDRLSLVLPAPQKLOQHTHEKPCMTSFSYVASAIPSD 492  
246 PEDILHLPEVLEHDFGSKLIVPKDRLSLVLPAPQKLOQHTHEKPCMTSFSYVASAIPSD 305  
493 LYFGSFCPGSGIKQIOVKONISVTLRTFAPSFQOASROGLTVSFIPIYKKEGVFTVTPD 552  
306 LYFGSFCPGSGIKQIOVKONISVTLRTFAPSFQOASROGLTVSFIPIYKKEGVFTVTPD 365  
553 TKSATYVLRTPNMDRGLPSLTSVSNISVPRDOVACITFEKERSGVVCCOTGRAFMIIQEQR 612  
366 TKSATYVLRTPNMDRGLPSLTSVSNISVPRDOVACITFEKERSGVVCCOTGRAFMIIQEQR 425  
613 TRAEIIFSLDEVDLKPSPFHHSFMVNISNCSPTSGKQDLIFSVTLPRVTDLVIIILA 672  
426 TRAEIIFSLDEVDLKPSPFHHSFMVNISNCSPTSGKQDLIFSVTLPRVTDLVIIILA 485  
673 AVGGGVLLLSALGLIICVKKKKKKTKNGPAVGIYNGNINTEMPROPKKFQGRKDNDSH 732  
486 AVGGGVLLLSALGLIICVKKKKKKTKNGPAVGIYNGNINTEMPROPKKFQGRKDNDSH 545  
733 VYAVIEDTMVYGHLLQDSSGSFLOPEVDVTRPFGQMGVCPSPPTIGSRAPTAKLATEE 792  
546 VYAVIEDTMVYGHLLQDSSGSFLOPEVDVTRPFGQMGVCPSPPTIGSRAPTAKLATEE 605  
793 PPPRSPSESEPRYTFSPHNNGDVSKOTDIPLNTQEMEPAE 836  
606 PPPRSPSESEPRYTFSPHNNGDVSKOTDIPLNTQEMEPAE 649

RESULT 8

PCT-US01-43523-162  
Sequence 162, Application PC/TUS0143523  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: DeForge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerltzen, Mary E.  
APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zhen  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
TITLE OF INVENTION: ACIDS ENCODING THE SAME  
FILE REFERENCE: P3330R1C331  
CURRENT FILING DATE: 2002-10-02  
PRIOR APPLICATION NUMBER: 60/049911  
PRIOR FILING DATE: 1997-06-18  
PRIOR APPLICATION NUMBER: 60/056974  
PRIOR FILING DATE: 1997-08-26  
PRIOR APPLICATION NUMBER: 60/059113  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059115  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059117  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059122  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059184  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059263  
PRIOR FILING DATE: 1997-09-18  
PRIOR APPLICATION NUMBER: 60/059352  
PRIOR FILING DATE: 1997-09-19  
PRIOR APPLICATION NUMBER: 60/059588  
PRIOR FILING DATE: 1997-09-19  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO: 162  
LENGTH: 343  
TYPE: PRT  
ORGANISM: Homo Sapien  
PCT-US01-43523-162

Query Match 40.8%; Score 341; DB 1; Length 343;  
Best Local Similarity 100.0%; Pred. No. 36-304;  
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MAGLNGVSIALLGVLLGAARLPAGAEFIALPRESNITVLKLGPTLLAKRCYIVI 60  
1 MAGLNGVSIALLGVLLGAARLPAGAEFIALPRESNITVLKLGPTLLAKRCYIVI 60  
61 SKRHTMISTSGRIVTTFSCQSPENHFVLEIQKNIDCMGCPGFGVLOPSTSLPT 120  
61 SKRHTMISTSGRIVTTFSCQSPENHFVLEIQKNIDCMGCPGFGVLOPSTSLPT 120  
121 LNRFTIMVKAHKSIGLEQPSIRPLROIQGESCPDVTHSISGRIDATVIRIGTFCSN 180  
121 LNRFTIMVKAHKSIGLEQPSIRPLROIQGESCPDVTHSISGRIDATVIRIGTFCSN 180  
181 GTVSRIMQEGVKALHLPMFHPRVSGFSIANRSSIKRLCTIBSVFGECSATLMSANY 240  
181 GTVSRIMQEGVKALHLPMFHPRVSGFSIANRSSIKRLCTIBSVFGECSATLMSANY 240  
241 PEGFPEDELMTWQFVPAHLRASVFLNPNLNCERKEERYIIPGTTPEVFKLEDK 300  
241 PEGFPEDELMTWQFVPAHLRASVFLNPNLNCERKEERYIIPGTTPEVFKLEDK 300  
301 QPGMAGNPNLSLOGCDQDASPGILRLQFVLVQHPONES 341  
301 QPGMAGNPNLSLOGCDQDASPGILRLQFVLVQHPONES 341

RESULT 9

PCT-US02-24563-162

```

; Sequence 162, Application PC/TUS0224563
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C331
; CURRENT APPLICATION NUMBER: PCT/US02/24563
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 162
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Homo Sapien
PCT-US02-24563-162

```

```

Query Match      40.8%; Score 341; DB 1; Length 343;
Best Local Similarity 100.0%; Pred No. 3e-304;
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MAGLNGVSIALLGVLLGAAALPRGAFAEIALPRESNITVLKIGTPTLLAKPCYIV 60
DB 1 MAGLNGVSIALLGVLLGAAALPRGAFAEIALPRESNITVLKIGTPTLLAKPCYIV 60
QY 61 SKRHITMILSIKSERIVFTFSCSPENHVIETQKIDCMGSGCPGEVQLQPSISLPT 120
DB 61 SKRHITMILSIKSERIVFTFSCSPENHVIETQKIDCMGSGCPGEVQLQPSISLPT 120
QY 121 LNRFTFMDVKAKHSIGLELOFSIPRLRQIGPESCPDGVTHISGRIDATVIRIGFCGN 180
DB 121 LNRFTFMDVKAKHSIGLELOFSIPRLRQIGPESCPDGVTHISGRIDATVIRIGFCGN 180
QY 181 GTVSRIKMOEGVSMALHLPFHRNVSQFSINRSSIKRLCTIESEYFEEGSAITMSANY 240
DB 181 GTVSRIKMOEGVSMALHLPFHRNVSQFSINRSSIKRLCTIESEYFEEGSAITMSANY 240
QY 241 PGGFPEDELMTWQFVPAHLRASVSFLNPNLSNCKEKREVEYYIPGSTTNPBEVFKLEDK 300

```

```

DB 241 PGGFPEDELMTWQFVPAHLRASVSFLNPNLSNCKEKREVEYYIPGSTTNPBEVFKLEDK 300
QY 301 QPGNMGNFNLSLQCCDDAOSPGILRLQFQVLYVHPONES 341
DB 301 QPGNMGNFNLSLQCCDDAOSPGILRLQFQVLYVHPONES 341

RESULT 10
US-10-028-072-162
; Sequence 162, Application US/10028072
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION:
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/028,072
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059836
; PRIOR FILING DATE: 1997-09-24
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062285
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062814
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/062816
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063045
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063082
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/063127
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063327
; PRIOR FILING DATE: 1997-10-27

```

PRIOR APPLICATION NUMBER: 60/063329  
 PRIOR FILING DATE: 1997-10-27  
 PRIOR APPLICATION NUMBER: 60/063350  
 PRIOR FILING DATE: 1997-10-28  
 PRIOR APPLICATION NUMBER: 60/063561  
 PRIOR FILING DATE: 1997-10-28  
 PRIOR APPLICATION NUMBER: 60/063704  
 PRIOR FILING DATE: 1997-10-29  
 PRIOR APPLICATION NUMBER: 60/063733  
 PRIOR FILING DATE: 1997-10-29  
 PRIOR APPLICATION NUMBER: 60/063735  
 PRIOR FILING DATE: 1997-10-29  
 PRIOR APPLICATION NUMBER: 60/063738  
 PRIOR FILING DATE: 1997-10-29  
 PRIOR APPLICATION NUMBER: 60/063755  
 PRIOR FILING DATE: 1997-10-17  
 PRIOR APPLICATION NUMBER: 60/064248  
 PRIOR FILING DATE: 1997-11-03  
 PRIOR APPLICATION NUMBER: 60/064809  
 PRIOR FILING DATE: 1997-11-07  
 PRIOR APPLICATION NUMBER: 60/065186  
 PRIOR FILING DATE: 1997-11-12  
 PRIOR APPLICATION NUMBER: 60/065846  
 PRIOR FILING DATE: 1997-11-17  
 PRIOR APPLICATION NUMBER: 60/066364  
 PRIOR FILING DATE: 1997-11-21  
 PRIOR APPLICATION NUMBER: 60/066453  
 PRIOR FILING DATE: 1997-11-24  
 PRIOR APPLICATION NUMBER: 60/066511  
 PRIOR FILING DATE: 1997-11-24  
 PRIOR APPLICATION NUMBER: 60/066770  
 PRIOR FILING DATE: 1997-11-24  
 PRIOR APPLICATION NUMBER: 60/069212  
 PRIOR FILING DATE: 1997-12-11  
 PRIOR APPLICATION NUMBER: 60/069278  
 PRIOR FILING DATE: 1997-12-11  
 PRIOR APPLICATION NUMBER: 60/069334  
 PRIOR FILING DATE: 1997-12-11  
 PRIOR APPLICATION NUMBER: 60/069694  
 PRIOR FILING DATE: 1997-12-16  
 PRIOR APPLICATION NUMBER: 60/072320  
 PRIOR FILING DATE: 1998-01-23  
 PRIOR APPLICATION NUMBER: 60/073612  
 PRIOR FILING DATE: 1998-02-04  
 PRIOR APPLICATION NUMBER: 60/074086  
 PRIOR FILING DATE: 1998-02-09  
 PRIOR APPLICATION NUMBER: 60/074092  
 PRIOR FILING DATE: 1998-02-09  
 PRIOR APPLICATION NUMBER: 60/077791  
 PRIOR FILING DATE: 1998-03-12  
 PRIOR APPLICATION NUMBER: 60/078910  
 PRIOR FILING DATE: 1998-03-20  
 PRIOR APPLICATION NUMBER: 60/079294  
 PRIOR FILING DATE: 1998-03-25  
 PRIOR APPLICATION NUMBER: 60/079663  
 PRIOR FILING DATE: 1998-02-27  
 PRIOR APPLICATION NUMBER: 60/079728  
 PRIOR FILING DATE: 1998-03-27  
 PRIOR APPLICATION NUMBER: 60/080165  
 PRIOR FILING DATE: 1998-03-31  
 PRIOR APPLICATION NUMBER: 60/081203  
 PRIOR FILING DATE: 1998-04-09  
 PRIOR APPLICATION NUMBER: 60/081229  
 PRIOR FILING DATE: 1998-04-09  
 PRIOR APPLICATION NUMBER: 60/081695  
 PRIOR FILING DATE: 1998-04-14  
 PRIOR APPLICATION NUMBER: 60/081817  
 PRIOR FILING DATE: 1998-04-15  
 PRIOR APPLICATION NUMBER: 60/081818  
 PRIOR FILING DATE: 1998-04-15  
 PRIOR APPLICATION NUMBER: 60/082999  
 PRIOR FILING DATE: 1998-04-24  
 PRIOR APPLICATION NUMBER: 60/083322

PRIOR FILING DATE: 1998-04-28  
 PRIOR APPLICATION NUMBER: 60/083545  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/084600  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/084627  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/084637  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/085149  
 PRIOR FILING DATE: 1998-05-12  
 PRIOR APPLICATION NUMBER: 60/085323  
 PRIOR FILING DATE: 1998-05-13  
 PRIOR APPLICATION NUMBER: 60/085338  
 PRIOR FILING DATE: 1998-05-13  
 PRIOR APPLICATION NUMBER: 60/085339  
 PRIOR FILING DATE: 1998-05-13  
 PRIOR APPLICATION NUMBER: 60/085579  
 PRIOR FILING DATE: 1998-05-15  
 PRIOR APPLICATION NUMBER: 60/085697  
 PRIOR FILING DATE: 1998-05-15  
 PRIOR APPLICATION NUMBER: 60/085704  
 PRIOR FILING DATE: 1998-05-15  
 PRIOR APPLICATION NUMBER: 60/086414  
 PRIOR FILING DATE: 1998-05-22  
 PRIOR APPLICATION NUMBER: 60/086430  
 PRIOR FILING DATE: 1998-05-22  
 PRIOR APPLICATION NUMBER: 60/087106  
 PRIOR FILING DATE: 1998-05-28  
 PRIOR APPLICATION NUMBER: 60/088026  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088730  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088741  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088810  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088858  
 PRIOR FILING DATE: 1998-06-11  
 PRIOR APPLICATION NUMBER: 60/089532  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089599  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089907  
 PRIOR FILING DATE: 1998-06-18  
 PRIOR APPLICATION NUMBER: 60/089947  
 PRIOR FILING DATE: 1998-06-19  
 PRIOR APPLICATION NUMBER: 60/090349  
 PRIOR FILING DATE: 1998-06-23  
 PRIOR APPLICATION NUMBER: 60/090429  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090445  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090538  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090863  
 PRIOR FILING DATE: 1998-06-26  
 PRIOR APPLICATION NUMBER: 60/091360  
 PRIOR FILING DATE: 1998-07-01  
 PRIOR APPLICATION NUMBER: 60/091519  
 PRIOR FILING DATE: 1998-07-02  
 PRIOR APPLICATION NUMBER: 60/091982  
 PRIOR FILING DATE: 1998-07-07  
 PRIOR APPLICATION NUMBER: 60/093339

Query Match 40.8%; Score 341; DB 26; Length 343;  
 Best Local Similarity 100.0%; Pred. No. 3e-304;  
 Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGLNCGVSIALLGVLLGAARLPGAAAFRIALPRESNITVLLKGTPTLLAKPCYIVI 60  
 DB 1 MAGLNCGVSIALLGVLLGAARLPGAAAFRIALPRESNITVLLKGTPTLLAKPCYIVI 60

QY 61 SKRHITMSTIKSGERIVTFPSCSPENHVFIEIQKIDCMSCGPCPGEVQLOPSTSLPT 120  
DB 61 SKRHITMSTIKSGERIVTFPSCSPENHVFIEIQKIDCMSCGPCPGEVQLOPSTSLPT 120  
QY 121 LNRFTIWDVKAHKSIGLEIQFSIPRLROIQPGESCPDGVTHSISGRIDATVVRIGTFCGN 180  
DB 121 LNRFTIWDVKAHKSIGLEIQFSIPRLROIQPGESCPDGVTHSISGRIDATVVRIGTFCGN 180  
QY 181 GTVSRIKMOEGVMALHLPWFHPRNVSGFSIANRSSIKRLCTIESEVFEBSGATLMSANY 240  
DB 181 GTVSRIKMOEGVMALHLPWFHPRNVSGFSIANRSSIKRLCTIESEVFEBSGATLMSANY 240  
QY 241 PEGFPEDELMTQFVVPVPAHLRASVSFLNPNLSNCRKEERVEYYIPGSTTNEVEVFLBCK 300  
DB 241 PEGFPEDELMTQFVVPVPAHLRASVSFLNPNLSNCRKEERVEYYIPGSTTNEVEVFLBCK 300  
QY 301 QPGNMGNFNLSLQCGDDAOSPGILRLQFOVLVQHPONES 341  
DB 301 QPGNMGNFNLSLQCGDDAOSPGILRLQFOVLVQHPONES 341

## RESULT 11

US-10-121-040-162  
Sequence 162, Application US/10121040

## GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: DeForge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3330R1C7  
CURRENT APPLICATION NUMBER: US/10/121,040  
CURRENT FILING DATE: 2002-04-11  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 162  
LENGTH: 343  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-10-121-040-162

Query Match 40.8%; Score 341; DB 27; Length 343;

Best Local Similarity 100.0%; Pred. No. 3e-304;  
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGLNCGVSIALLGVLLGAARLPFGAEAFETALPRESNITVLIKGTPTLLAKPCYIYI 60  
DB 1 MAGLNCGVSIALLGVLLGAARLPFGAEAFETALPRESNITVLIKGTPTLLAKPCYIYI 60  
QY 61 SKRHITMSTIKSGERIVTFPSCSPENHVFIEIQKIDCMSCGPCPGEVQLOPSTSLPT 120  
DB 61 SKRHITMSTIKSGERIVTFPSCSPENHVFIEIQKIDCMSCGPCPGEVQLOPSTSLPT 120  
QY 121 LNRFTIWDVKAHKSIGLEIQFSIPRLROIQPGESCPDGVTHSISGRIDATVVRIGTFCGN 180  
DB 121 LNRFTIWDVKAHKSIGLEIQFSIPRLROIQPGESCPDGVTHSISGRIDATVVRIGTFCGN 180  
QY 181 GTVSRIKMOEGVMALHLPWFHPRNVSGFSIANRSSIKRLCTIESEVFEBSGATLMSANY 240  
DB 181 GTVSRIKMOEGVMALHLPWFHPRNVSGFSIANRSSIKRLCTIESEVFEBSGATLMSANY 240

DB 181 GTVSRIKMOEGVMALHLPWFHPRNVSGFSIANRSSIKRLCTIESEVFEBSGATLMSANY 240  
QY 241 PEGFPEDELMTQFVVPVPAHLRASVSFLNPNLSNCRKEERVEYYIPGSTTNEVEVFLBCK 300  
DB 241 PEGFPEDELMTQFVVPVPAHLRASVSFLNPNLSNCRKEERVEYYIPGSTTNEVEVFLBCK 300  
QY 301 QPGNMGNFNLSLQCGDDAOSPGILRLQFOVLVQHPONES 341  
DB 301 QPGNMGNFNLSLQCGDDAOSPGILRLQFOVLVQHPONES 341

## RESULT 12

US-10-121-041-162  
Sequence 162, Application US/10121041

## GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: DeForge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3330R1C9  
CURRENT APPLICATION NUMBER: US/10/121,041  
CURRENT FILING DATE: 2002-04-11  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 162  
LENGTH: 343  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-10-121-041-162

Query Match 40.8%; Score 341; DB 27; Length 343;

Best Local Similarity 100.0%; Pred. No. 3e-304;  
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGLNCGVSIALLGVLLGAARLPFGAEAFETALPRESNITVLIKGTPTLLAKPCYIYI 60  
DB 1 MAGLNCGVSIALLGVLLGAARLPFGAEAFETALPRESNITVLIKGTPTLLAKPCYIYI 60  
QY 61 SKRHITMSTIKSGERIVTFPSCSPENHVFIEIQKIDCMSCGPCPGEVQLOPSTSLPT 120  
DB 61 SKRHITMSTIKSGERIVTFPSCSPENHVFIEIQKIDCMSCGPCPGEVQLOPSTSLPT 120  
QY 121 LNRFTIWDVKAHKSIGLEIQFSIPRLROIQPGESCPDGVTHSISGRIDATVVRIGTFCGN 180  
DB 121 LNRFTIWDVKAHKSIGLEIQFSIPRLROIQPGESCPDGVTHSISGRIDATVVRIGTFCGN 180  
QY 181 GTVSRIKMOEGVMALHLPWFHPRNVSGFSIANRSSIKRLCTIESEVFEBSGATLMSANY 240  
DB 181 GTVSRIKMOEGVMALHLPWFHPRNVSGFSIANRSSIKRLCTIESEVFEBSGATLMSANY 240  
QY 241 PEGFPEDELMTQFVVPVPAHLRASVSFLNPNLSNCRKEERVEYYIPGSTTNEVEVFLBCK 300  
DB 241 PEGFPEDELMTQFVVPVPAHLRASVSFLNPNLSNCRKEERVEYYIPGSTTNEVEVFLBCK 300  
QY 301 QPGNMGNFNLSLQCGDDAOSPGILRLQFOVLVQHPONES 341  
DB 301 QPGNMGNFNLSLQCGDDAOSPGILRLQFOVLVQHPONES 341

## RESULT 13

US-10-121-042-162

; Sequence 162, Application US/10121042

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tamas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; TITLE OF INVENTION: ACIDS ENCODING THE SAME  
; FILE REFERENCE: P3330R1C10  
; CURRENT APPLICATION NUMBER: US/10/121,042  
; CURRENT FILING DATE: 2002-04-11  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 162  
; LENGTH: 343  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-121-042-162

Query Match 40.8%; Score 341; DB 27; Length 343;  
Best Local Similarity 100.0%; Pred. No. 3e-304;  
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGNCGVSIALGVLLGAARLPRGAEAFETALPRESNITVLKLGPTTLAKPCYVI 60  
DB 1 MAGNCGVSIALGVLLGAARLPRGAEAFETALPRESNITVLKLGPTTLAKPCYVI 60  
QY 61 SKRHITMSTKSGERIVTFSCQSPENHFVIEIQKNDMSGPCPGFGEVQLOPSTSLPT 120  
DB 61 SKRHITMSTKSGERIVTFSCQSPENHFVIEIQKNDMSGPCPGFGEVQLOPSTSLPT 120  
QY 121 LNRFTIMVVKAKHSIGLELQFSIPRLRQIGPESCPDGVTHSISGRIDATVVRIGTFPCSN 180  
DB 121 LNRFTIMVVKAKHSIGLELQFSIPRLRQIGPESCPDGVTHSISGRIDATVVRIGTFPCSN 180  
QY 181 GTVSRIRKQEGVKALHLPMFPHRVNVSFSTANRSSIRKLCIIESVFEGESATLMSANY 240  
DB 181 GTVSRIRKQEGVKALHLPMFPHRVNVSFSTANRSSIRKLCIIESVFEGESATLMSANY 240  
QY 241 PEGFPEDELMTQFVPAHLRASVSFLNPNLSNCRKEKREVEYYIPGSTTNPVEVKLEDX 300  
DB 241 PEGFPEDELMTQFVPAHLRASVSFLNPNLSNCRKEKREVEYYIPGSTTNPVEVKLEDX 300  
QY 301 QPGNAGNPNLSLQCCDDAOSPGILRLQFQVLYVHPONES 341  
DB 301 QPGNAGNPNLSLQCCDDAOSPGILRLQFQVLYVHPONES 341

## RESULT 14

US-10-121-043-162

; Sequence 162, Application US/10121043

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen

; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tamas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; TITLE OF INVENTION: ACIDS ENCODING THE SAME  
; FILE REFERENCE: P3330R1C15  
; CURRENT APPLICATION NUMBER: US/10/121,043  
; CURRENT FILING DATE: 2002-04-12  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 162  
; LENGTH: 343  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-121-043-162

Query Match 40.8%; Score 341; DB 27; Length 343;  
Best Local Similarity 100.0%; Pred. No. 3e-304;  
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGNCGVSIALGVLLGAARLPRGAEAFETALPRESNITVLKLGPTTLAKPCYVI 60  
DB 1 MAGNCGVSIALGVLLGAARLPRGAEAFETALPRESNITVLKLGPTTLAKPCYVI 60  
QY 61 SKRHITMSTKSGERIVTFSCQSPENHFVIEIQKNDMSGPCPGFGEVQLOPSTSLPT 120  
DB 61 SKRHITMSTKSGERIVTFSCQSPENHFVIEIQKNDMSGPCPGFGEVQLOPSTSLPT 120  
QY 121 LNRFTIMVVKAKHSIGLELQFSIPRLRQIGPESCPDGVTHSISGRIDATVVRIGTFPCSN 180  
DB 121 LNRFTIMVVKAKHSIGLELQFSIPRLRQIGPESCPDGVTHSISGRIDATVVRIGTFPCSN 180  
QY 181 GTVSRIRKQEGVKALHLPMFPHRVNVSFSTANRSSIRKLCIIESVFEGESATLMSANY 240  
DB 181 GTVSRIRKQEGVKALHLPMFPHRVNVSFSTANRSSIRKLCIIESVFEGESATLMSANY 240  
QY 241 PEGFPEDELMTQFVPAHLRASVSFLNPNLSNCRKEKREVEYYIPGSTTNPVEVKLEDX 300  
DB 241 PEGFPEDELMTQFVPAHLRASVSFLNPNLSNCRKEKREVEYYIPGSTTNPVEVKLEDX 300  
QY 301 QPGNAGNPNLSLQCCDDAOSPGILRLQFQVLYVHPONES 341  
DB 301 QPGNAGNPNLSLQCCDDAOSPGILRLQFQVLYVHPONES 341

## RESULT 15

US-10-121-044-162

; Sequence 162, Application US/10121044

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tamas, Daniel  
; APPLICANT: Watanabe, Colin K

```

; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C5
; CURRENT APPLICATION NUMBER: US/10/121,044
; PRIORITY DATE: 2002-04-11
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 162
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-10-121-044-162

```

```

Query Match      40.8%; Score 341; DB 27; Length 343;
Best Local Similarity 100.0%; Pred. No. 3e-304;
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 MAGLNGVSIALIGVLLGAARLPAGAAFEIALPRESNITVLIKLGPTLLAKPCYIVI 60
Db      1 MAGLNGVSIALIGVLLGAARLPAGAAFEIALPRESNITVLIKLGPTLLAKPCYIVI 60
QY      61 SKRHITMLSIKSGERTVFTFSCQSPENHFVIEIQKNIDCMGSPCPGEVQLQPTSLPT 120
Db      61 SKRHITMLSIKSGERTVFTFSCQSPENHFVIEIQKNIDCMGSPCPGEVQLQPTSLPT 120
QY      121 LNRFTIMDVKAHKSIGLELOFSIPRLQIQGESCPDGVTHSISGRIDATVVRIGTFCSN 180
Db      121 LNRFTIMDVKAHKSIGLELOFSIPRLQIQGESCPDGVTHSISGRIDATVVRIGTFCSN 180
QY      181 GTVSRIKMQEGVMALHLPWFHPRNVSGFSIANRSSIKRLCTIESVFESEGSATLMSANY 240
Db      181 GTVSRIKMQEGVMALHLPWFHPRNVSGFSIANRSSIKRLCTIESVFESEGSATLMSANY 240
QY      241 PEGFPDELTMTQFVVPALHRAVSFLNPLNSCERKEKREVEYYIPGSTNPVEVFKLEDK 300
Db      241 PEGFPDELTMTQFVVPALHRAVSFLNPLNSCERKEKREVEYYIPGSTNPVEVFKLEDK 300
QY      301 QGNNMAGNFNLISLOGCDQDAQSPGILRLQFOVLVQHPONES 341
Db      301 QGNNMAGNFNLISLOGCDQDAQSPGILRLQFOVLVQHPONES 341

```

Search completed: February 20, 2004, 19:59:03  
Job time : 363 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OW protein - protein search, using SW model

Run on: February 20, 2004, 19:49:29 ; Search time 44 Seconds  
(without alignments)  
1495.299 Million cell updates/sec

Title: US-09-899-569A-4  
Perfect score: 836  
Sequence: 1 MAGNCGVSIALLGVLLGA.....SSKDTDIPLNTQEBEPBAE 836

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 378584 seqs, 7869973 residues

Word size : 0

Total number of hits satisfying chosen parameters: 378584

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Pending\_Patents\_AA\_New:\*

1: /cgn2\_6/ptodata/1/paa/PCT\_NEW\_COMB.pep:\*\n2: /cgn2\_6/ptodata/1/paa/US06\_NEW\_COMB.pep:\*\n3: /cgn2\_6/ptodata/1/paa/US07\_NEW\_COMB.pep:\*\n4: /cgn2\_6/ptodata/1/paa/US08\_NEW\_COMB.pep:\*\n5: /cgn2\_6/ptodata/1/paa/US09\_NEW\_COMB.pep:\*\n6: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep:\*\n7: /cgn2\_6/ptodata/1/paa/US60\_NEW\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9	1.1	293	6	US-10-481-265-105
2	8	1.0	134	1	PCT-US03-38896-178
3	8	1.0	147	6	US-10-669-657-162
4	8	1.0	187	1	PCT-US03-38829-199
5	8	1.0	193	1	PCT-US03-38896-177
6	8	1.0	221	5	US-09-978-360A-436
7	8	1.0	221	6	US-10-679-063-6997
8	8	1.0	221	6	US-10-679-063-8684
9	8	1.0	221	6	US-10-679-063-11202
10	8	1.0	222	6	US-10-425-114A-50595
11	8	1.0	276	6	US-10-425-114A-11744
12	8	1.0	348	6	US-10-739-930-8397
13	8	1.0	377	1	PCT-US03-19153-296
14	8	1.0	377	6	US-10-463-720-296
15	8	1.0	485	6	US-10-425-114A-65537
16	8	1.0	487	6	US-10-425-114A-64916
17	8	1.0	499	6	US-10-425-114A-40938
18	8	1.0	499	6	US-10-425-114A-42517
19	8	1.0	631	5	US-09-614-150A-26727
20	8	1.0	634	6	US-10-425-114A-59648
21	7	0.8	19	5	US-09-341-590A-122
22	7	0.8	30	6	US-10-296-734-714
23	7	0.8	30	6	US-10-296-734-716
24	7	0.8	61	6	US-10-679-063-24192
25	7	0.8	75	6	US-10-705-531-4
26	7	0.8	77	5	US-09-796-692A-1805

27	7	0.8	78	6	US-10-415-182A-10176	Sequence 10176, A
28	7	0.8	79	6	US-10-415-182A-2392	Sequence 2392, Ap
29	7	0.8	82	6	US-10-425-114A-45433	Sequence 45433, A
30	7	0.8	82	6	US-10-462-601-29	Sequence 29, Appl
31	7	0.8	83	5	US-09-614-150A-32895	Sequence 32895, A
32	7	0.8	90	6	US-10-425-114A-58082	Sequence 58082, A
33	7	0.8	97	6	US-10-460-594-94	Sequence 94, Appl
34	7	0.8	113	6	US-10-679-063-18469	Sequence 18469, A
35	7	0.8	129	6	US-10-425-114A-67250	Sequence 67250, A
36	7	0.8	140	6	US-10-425-114A-58799	Sequence 58799, A
37	7	0.8	146	6	US-10-679-063-5291	Sequence 5291, Ap
38	7	0.8	157	6	US-10-425-114A-63652	Sequence 63652, Ap
39	7	0.8	165	6	US-10-679-063-4799	Sequence 4799, Ap
40	7	0.8	169	6	US-10-679-063-5043	Sequence 5043, Ap
41	7	0.8	176	6	US-10-679-063-4736	Sequence 4736, Ap
42	7	0.8	177	6	US-10-425-114A-46400	Sequence 46400, A
43	7	0.8	177	6	US-10-425-114A-73004	Sequence 73004, A
44	7	0.8	179	6	US-10-425-114A-49322	Sequence 49322, A
45	7	0.8	185	6	US-10-425-114A-53476	Sequence 53476, A

## ALIGNMENTS

RESULT 1  
US-10-481-265-105  
Sequence 105: Application US/10481265  
GENERAL INFORMATION:  
APPLICANT: James, Brian William  
APPLICANT: Bacon, Joanna  
TITLE OF INVENTION: Mycobacterial Antigens Expressed Under Low Oxygen Tension  
FILE REFERENCE: 1581.1020000  
CURRENT APPLICATION NUMBER: US/10/481,265  
PRIOR FILING DATE: 2003-12-19  
PRIOR APPLICATION NUMBER: GB 0115365.9  
PRIOR FILING DATE: 2001-06-22  
PRIOR APPLICATION NUMBER: GB 0121780.1  
PRIOR FILING DATE: 2001-09-07  
PRIOR APPLICATION NUMBER: PCT/GB02/02845  
PRIOR FILING DATE: 2002-06-21  
NUMBER OF SEQ ID NOS: 138  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 105  
LENGTH: 293  
TYPE: PRT  
ORGANISM: Mycobacterium tuberculosis  
US-10-481-265-105  
Query Match 1.1%; Score 9; DB 6; Length 293;  
Best Local Similarity 100.0%; Pred. No. 8.2;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
DB 674 VGGGVLLLS 682  
67 VGGGVLLLS 75  
RESULT 2  
PCT-US03-38896-178  
Sequence 178: Application PC/TUS0338896  
GENERAL INFORMATION:  
APPLICANT: diadexue, Inc.  
APPLICANT: Macina, Roberto  
APPLICANT: Turner, Leah  
APPLICANT: Sun, Yongming  
TITLE OF INVENTION: Compositions, Splice Variants and Methods Relating to Lung Specific  
FILE REFERENCE: DEX-0454  
CURRENT APPLICATION NUMBER: PCT/US03/38896  
CURRENT FILING DATE: 2003-12-08  
PRIOR APPLICATION NUMBER: US 60/431,516

; PRIOR FILING DATE: 2002-12-06  
 ; PRIOR APPLICATION NUMBER: US 60/431,510  
 ; PRIOR FILING DATE: 2002-12-06  
 ; PRIOR APPLICATION NUMBER: US 60/431,307  
 ; PRIOR FILING DATE: 2002-12-06  
 ; NUMBER OF SEQ ID NOS: 279  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 178  
 ; LENGTH: 134  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapien  
 PCT-US03-38896-178

Query Match 1.0%; Score 8; DB 1; Length 134;  
 Best Local Similarity 100.0%; Pred. No. 36;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 693 KKKKTKNK 700  
 DB 70 KKKKTKNK 77

RESULT 3  
 US-10-669-657-162  
 ; Sequence 162, Application US/10669657  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rosen, et. al  
 ; TITLE OF INVENTION: 124 Human Secreted Proteins  
 ; FILE REFERENCE: PS956  
 ; CURRENT APPLICATION NUMBER: US/10/669,657  
 ; CURRENT FILING DATE: 2003-09-25  
 ; PRIOR APPLICATION NUMBER: PCT/US02/09135  
 ; PRIOR FILING DATE: 2002-03-26  
 ; PRIOR APPLICATION NUMBER: 10/105,299  
 ; PRIOR FILING DATE: 2002-03-26  
 ; PRIOR APPLICATION NUMBER: 60/278,650  
 ; PRIOR FILING DATE: 2001-03-27  
 ; PRIOR APPLICATION NUMBER: 09/950,082  
 ; PRIOR FILING DATE: 2001-09-12  
 ; PRIOR APPLICATION NUMBER: 09/950,083  
 ; PRIOR FILING DATE: 2001-09-12  
 ; PRIOR APPLICATION NUMBER: 09/833,245  
 ; PRIOR FILING DATE: 2001-04-12  
 ; PRIOR APPLICATION NUMBER: PCT/US01/11988  
 ; PRIOR FILING DATE: 2001-04-12  
 ; PRIOR APPLICATION NUMBER: 60/331,287  
 ; PRIOR FILING DATE: 2001-11-13  
 ; PRIOR APPLICATION NUMBER: 60/277,340  
 ; PRIOR FILING DATE: 2001-03-21  
 ; PRIOR APPLICATION NUMBER: 60/306,171  
 ; PRIOR FILING DATE: 2001-07-19  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 493  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 162  
 ; LENGTH: 147  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-669-657-162

Query Match 1.0%; Score 8; DB 6; Length 147;  
 Best Local Similarity 100.0%; Pred. No. 38;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 LIGVLLG 19  
 DB 30 LIGVLLG 37

RESULT 4  
 PCT-US03-38829-199  
 ; Sequence 199, Application PC/TUS0338829  
 ; GENERAL INFORMATION:

; APPLICANT: diadexus, Inc.  
 ; APPLICANT: Macina, Roberto  
 ; APPLICANT: Turner, Leah  
 ; APPLICANT: Sun, Yongming  
 ; TITLE OF INVENTION: Compositions, Splice Variants and Methods Relating to Breast Spec  
 ; FILE REFERENCE: DEX-0453  
 ; CURRENT APPLICATION NUMBER: PCT/US03/38829  
 ; CURRENT FILING DATE: 2003-12-05  
 ; PRIOR APPLICATION NUMBER: US 60/431,145  
 ; PRIOR FILING DATE: 2002-12-05  
 ; NUMBER OF SEQ ID NOS: 253  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 199  
 ; LENGTH: 187  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapien  
 PCT-US03-38829-199

Query Match 1.0%; Score 8; DB 1; Length 187;  
 Best Local Similarity 100.0%; Pred. No. 44;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 LIGVLLG 19  
 DB 70 LIGVLLG 77

RESULT 5  
 PCT-US03-38896-177  
 ; Sequence 177, Application PC/TUS0338896  
 ; GENERAL INFORMATION:  
 ; APPLICANT: diadexus, Inc.  
 ; APPLICANT: Macina, Roberto  
 ; APPLICANT: Turner, Leah  
 ; APPLICANT: Sun, Yongming  
 ; APPLICANT: Liu, Shu-Hui  
 ; TITLE OF INVENTION: Compositions, Splice Variants and Methods Relating to Lung Specifi  
 ; FILE REFERENCE: DEX-0454  
 ; CURRENT APPLICATION NUMBER: PCT/US03/38896  
 ; CURRENT FILING DATE: 2003-12-08  
 ; PRIOR APPLICATION NUMBER: US 60/431,516  
 ; PRIOR FILING DATE: 2002-12-06  
 ; PRIOR APPLICATION NUMBER: US 60/431,510  
 ; PRIOR FILING DATE: 2002-12-06  
 ; PRIOR APPLICATION NUMBER: 60/431,307  
 ; PRIOR FILING DATE: 2002-12-06  
 ; NUMBER OF SEQ ID NOS: 279  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 177  
 ; LENGTH: 193  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapien  
 PCT-US03-38896-177

Query Match 1.0%; Score 8; DB 1; Length 193;  
 Best Local Similarity 100.0%; Pred. No. 45;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 693 KKKKTKNK 700  
 DB 68 KKKKTKNK 75

RESULT 6  
 US-09-978-360A-436  
 ; Sequence 436, Application US/09978360A  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Edwards, Jean-Baptiste Dumas Milne  
 ; APPLICANT: Duclert, Aymeric  
 ; APPLICANT: Bougueteloret, Lydie  
 ; APPLICANT: Jobert, Severin

APPLICANT: Clusel, Catherine  
; TITLE OF INVENTION: Complementary DNA's Encoding Proteins with Signal Peptides  
; FILE REFERENCE: 56, US4, CIP  
; CURRENT APPLICATION NUMBER: US/09/978,360A  
; CURRENT FILING DATE: 2003-10-15  
; PRIOR APPLICATION NUMBER: US 60/066,677  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: US 60/069,957  
; PRIOR FILING DATE: 1997-12-17  
; PRIOR APPLICATION NUMBER: US 60/074,121  
; PRIOR FILING DATE: 1998-02-09  
; PRIOR APPLICATION NUMBER: US 60/081,563  
; PRIOR FILING DATE: 1998-04-13  
; PRIOR APPLICATION NUMBER: US 60/096,116  
; PRIOR FILING DATE: 1998-08-10  
; PRIOR APPLICATION NUMBER: US 60/099,273  
; PRIOR FILING DATE: -09-04  
; PRIOR APPLICATION NUMBER: US 09/191,997  
; PRIOR FILING DATE: 1998-11-13  
; PRIOR APPLICATION NUMBER: US 09/215,435  
; PRIOR FILING DATE: 1998-12-17  
; PRIOR APPLICATION NUMBER: PCT/1998/02122  
; PRIOR FILING DATE: 1998-12-17  
; PRIOR APPLICATION NUMBER: US 09/247,155  
; PRIOR FILING DATE: 1999-02-09  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 810  
; SOFTWARE: Patent.pm  
; SEQ ID NO: 436  
; LENGTH: 221  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
; FEATURE:  
; NAME/KEY: SIGNAL  
; LOCATION: -28...-1  
; US-09-978-360A-436

Query Match 1.0%; Score 8; DB 5; Length 221;  
Best Local Similarity 100.0%; Pred. No. 49;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 LIGVLLLG 19  
Db 104 LIGVLLLG 111

RESULT 7  
US-10-679-063-6997  
; Sequence 6997, Application US/10679063  
; GENERAL INFORMATION:  
; APPLICANT: Edgerton, Michael D  
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES  
; FILE REFERENCE: 38-15(52054)B  
; CURRENT APPLICATION NUMBER: US/10/679,063  
; CURRENT FILING DATE: 2003-10-02  
; PRIOR APPLICATION NUMBER: 60/415,758  
; PRIOR FILING DATE: 2002-10-02  
; NUMBER OF SEQ ID NOS: 27373  
; SEQ ID NO: 6997  
; LENGTH: 221  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-10-679-063-6997

Query Match 1.0%; Score 8; DB 6; Length 221;  
Best Local Similarity 100.0%; Pred. No. 49;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 LIGVLLLG 19  
Db 104 LIGVLLLG 111

RESULT 8  
US-10-679-063-8684  
; Sequence 8684, Application US/10679063  
; GENERAL INFORMATION:  
; APPLICANT: Edgerton, Michael D  
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES  
; FILE REFERENCE: 38-15(52054)B  
; CURRENT APPLICATION NUMBER: US/10/679,063  
; CURRENT FILING DATE: 2003-10-02  
; PRIOR APPLICATION NUMBER: 60/415,758  
; PRIOR FILING DATE: 2002-10-02  
; NUMBER OF SEQ ID NOS: 27373  
; SEQ ID NO: 8684  
; LENGTH: 221  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-10-679-063-8684

Query Match 1.0%; Score 8; DB 6; Length 221;  
Best Local Similarity 100.0%; Pred. No. 49;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 LIGVLLLG 19  
Db 104 LIGVLLLG 111

RESULT 9  
US-10-679-063-11202  
; Sequence 11202, Application US/10679063  
; GENERAL INFORMATION:  
; APPLICANT: Edgerton, Michael D  
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES  
; FILE REFERENCE: 38-15(52054)B  
; CURRENT APPLICATION NUMBER: US/10/679,063  
; CURRENT FILING DATE: 2003-10-02  
; PRIOR APPLICATION NUMBER: 60/415,758  
; PRIOR FILING DATE: 2002-10-02  
; NUMBER OF SEQ ID NOS: 27373  
; SEQ ID NO: 11202  
; LENGTH: 221  
; TYPE: PRT  
; ORGANISM: Papio anubis  
; US-10-679-063-11202

Query Match 1.0%; Score 8; DB 6; Length 221;  
Best Local Similarity 100.0%; Pred. No. 49;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 LIGVLLLG 19  
Db 104 LIGVLLLG 111

RESULT 10  
US-10-425-114A-50595  
; Sequence 50595, Application US/10425114A  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114A  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO: 50595  
; LENGTH: 222  
; TYPE: PRT

ORGANISM: Zea mays  
FEATURE:  
OTHER INFORMATION: Clone ID: 700051071\_FLI.pep  
US-10-425-114A-50595

Query Match 1.0%; Score 8; DB 6; Length 222;  
Best Local Similarity 100.0%; Pred. No. 49;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 VLLGARN 22  
DB 210 VLLGARN 217

RESULT 11  
US-10-425-114A-71744  
Sequence 71744, Application US/10425114A  
GENERAL INFORMATION:  
APPLICANT: Liu, Jindong  
APPLICANT: Zhou, Yihua  
APPLICANT: Kovalic, David K.  
APPLICANT: Screen, Steven E  
APPLICANT: Tabaska, Jack E  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53313)B  
CURRENT APPLICATION NUMBER: US/10/425,114A  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 73128  
SEQ ID NO 71744  
LENGTH: 276  
TYPE: PRT  
ORGANISM: Zea mays  
FEATURE:  
OTHER INFORMATION: Clone ID: UC-ZMFLM017122B10\_FLI.pep  
US-10-425-114A-71744

Query Match 1.0%; Score 8; DB 6; Length 276;  
Best Local Similarity 100.0%; Pred. No. 56;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 773 PPSPTIC 780  
DB 74 PPSPTIC 81

RESULT 12  
US-10-739-930-8397  
Sequence 8397, Application US/10739930  
GENERAL INFORMATION:  
APPLICANT: Kovalic, David K.  
TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH  
TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT  
FILE REFERENCE: 38-21(53377)B  
CURRENT APPLICATION NUMBER: US/10/739,930  
CURRENT FILING DATE: 2003-12-18  
NUMBER OF SEQ ID NOS: 11088  
SEQ ID NO 8397  
LENGTH: 348  
TYPE: PRT  
ORGANISM: Zea mays  
FEATURE:  
OTHER INFORMATION: Clone ID: ZEAMA-23APR03-C6211\_2.p  
US-10-739-930-8397

Query Match 1.0%; Score 8; DB 6; Length 348;  
Best Local Similarity 100.0%; Pred. No. 64;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LUGARLP 24  
DB 197 LUGARLP 204

RESULT 13  
PCT-US03-19153-296  
Sequence 296, Application PC/TUS0319153  
GENERAL INFORMATION:  
APPLICANT: Diversa Corporation  
TITLE OF INVENTION: XYLANASES, NUCLEIC ACIDS ENCODING THEM AND  
TITLE OF INVENTION: METHODS FOR MAKING AND USING THEM  
FILE REFERENCE: 09010-290W01  
CURRENT APPLICATION NUMBER: PCT/US03/19153  
CURRENT FILING DATE: 2003-06-16  
PRIOR APPLICATION NUMBER: US 60/389,299  
PRIOR FILING DATE: 2002-06-14  
NUMBER OF SEQ ID NOS: 378  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 296  
LENGTH: 377  
TYPE: PRT  
ORGANISM: Unknown  
FEATURE:  
OTHER INFORMATION: Obtained from an environmental sample.  
NAME/KEY: SIGNAL  
LOCATION: (1)...(26)  
PCT-US03-19153-296

Query Match 1.0%; Score 8; DB 1; Length 377;  
Best Local Similarity 100.0%; Pred. No. 67;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 227 FEGBSAT 234  
DB 345 FEGBSAT 352

RESULT 14  
US-10-463-720-296  
Sequence 296, Application US/10463720  
GENERAL INFORMATION:  
APPLICANT: Steer, Brian  
APPLICANT: Callen, Walter  
APPLICANT: Healey, Shaun  
APPLICANT: Hazlewood, Geoff  
APPLICANT: Wu, Di  
APPLICANT: Blum, David  
APPLICANT: Esteghlalian, Alireza  
TITLE OF INVENTION: XYLANASES, NUCLEIC ACIDS ENCODING THEM AND  
TITLE OF INVENTION: METHODS FOR MAKING AND USING THEM  
FILE REFERENCE: 09010-290W01  
CURRENT APPLICATION NUMBER: US/10/463,720  
CURRENT FILING DATE: 2003-06-16  
PRIOR APPLICATION NUMBER: US 60/389,299  
PRIOR FILING DATE: 2002-06-14  
NUMBER OF SEQ ID NOS: 378  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 296  
LENGTH: 377  
TYPE: PRT  
ORGANISM: Unknown  
FEATURE:  
OTHER INFORMATION: Obtained from an environmental sample.  
NAME/KEY: SIGNAL  
LOCATION: (1)...(26)  
US-10-463-720-296

Query Match 1.0%; Score 8; DB 6; Length 377;  
Best Local Similarity 100.0%; Pred. No. 67;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 227 FEGBSAT 234  
DB 345 FEGBSAT 352

Db 345 FEGEGSAT 352

RESULT 15

US-10-425-114A-65537

Sequence 65537, Application US/10425114A

GENERAL INFORMATION:

APPLICANT: Liu, Jingdong

APPLICANT: Zhou, Yihua

APPLICANT: Kovalic, David K.

APPLICANT: Screen, Steven E

APPLICANT: Tabaska, Jack E

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21(5313)B

CURRENT APPLICATION NUMBER: US/10/425,114A

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 73128

SEQ ID NO 65537

LENGTH: 485

TYPE: PRT

ORGANISM: Zea mays

FEATURE:

OTHER INFORMATION: Clone ID: LIB4767-014-B10\_FLI.pcp

US-10-425-114A-65537

Query Match 1.0%; Score 8; DB 6; Length 485;

Best Local Similarity 100.0%; Pred. No. 78;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 653 LIFSCTLT 660

Db 107 LIFSCTLT 114

Search completed: February 20, 2004, 19:59:56.  
Job time : 46 secs

**This Page Blank (uspio)**